

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2004, 16:06:28 ; Search time 197 Seconds
(without alignments)
461.468 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158
Perfect score: 821
Sequence: 1 MKRVITLFAVLNGWSVNAW.....VAIKAGSLIAVLILRQTNNY 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 821 | 100.0 | 166 | 2 Q9F5R9 | Q9F5R9 escherichia |
| 2 | 821 | 100.0 | 300 | 1 FIMH_ECOLI | P08191 escherichia |
| 3 | 821 | 100.0 | 300 | 2 Q6JKV6 | Q6JKV6 escherichia |
| 4 | 821 | 100.0 | 300 | 2 Q84DW2 | Q84DW2 escherichia |
| 5 | 821 | 100.0 | 300 | 2 Q9F5S0 | Q9F5S0 escherichia |
| 6 | 817 | 99.5 | 300 | 2 Q7BYN8 | Q7BYN8 shigella fl |
| 7 | 817 | 99.5 | 300 | 2 Q83TY4 | Q83TY4 escherichia |
| 8 | 817 | 99.5 | 300 | 2 Q84DW3 | Q84DW3 escherichia |
| 9 | 817 | 99.5 | 300 | 2 Q83B73 | Q83B73 shigella fl |
| 10 | 815 | 99.3 | 300 | 2 Q9S494 | Q9S494 escherichia |
| 11 | 815 | 99.3 | 300 | 2 Q9S496 | Q9S496 escherichia |
| 12 | 815 | 99.3 | 300 | 2 Q9S497 | Q9S497 escherichia |
| 13 | 815 | 99.3 | 300 | 2 Q9R5Y2 | Q9R5Y2 escherichia |
| 14 | 814 | 99.1 | 300 | 2 Q84DW4 | Q84DW4 escherichia |
| 15 | 814 | 99.1 | 300 | 2 Q9F5S1 | Q9F5S1 escherichia |
| 16 | 813 | 99.0 | 300 | 2 Q6JKU4 | Q6JKU4 escherichia |
| 17 | 812 | 98.9 | 300 | 2 Q6JKU5 | Q6JKU5 escherichia |
| 18 | 812 | 98.9 | 300 | 2 Q6JKU6 | Q6JKU6 escherichia |
| 19 | 812 | 98.9 | 300 | 2 Q6JKV7 | Q6JKV7 escherichia |
| 20 | 811 | 98.8 | 300 | 2 Q6JKV8 | Q6JKV8 escherichia |
| 21 | 811 | 98.8 | 300 | 2 Q7ABM4 | Q7ABM4 escherichia |
| 22 | 811 | 98.8 | 300 | 2 Q8XBA6 | Q8XBA6 escherichia |
| 23 | 810 | 98.7 | 300 | 2 Q9S495 | Q9S495 escherichia |
| 24 | 809 | 98.5 | 300 | 2 Q6JKV8 | Q6JKV8 escherichia |
| 25 | 809 | 98.5 | 300 | 2 Q84DW0 | Q84DW0 escherichia |
| 26 | 809 | 98.5 | 300 | 2 Q9F6Z7 | Q9F6Z7 escherichia |
| 27 | 809 | 98.5 | 303 | 2 Q8CVF8 | Q8CVF8 escherichia |
| 28 | 808 | 98.4 | 300 | 2 Q84DW1 | Q84DW1 escherichia |
| 29 | 806 | 98.2 | 300 | 2 Q87634 | Q87634 escherichia |
| 30 | 806 | 98.2 | 303 | 2 Q9AP05 | Q9AP05 escherichia |
| 31 | 804 | 97.9 | 300 | 2 Q84DW5 | Q84DW5 escherichia |

| | | | | | |
|----|-------|------|-----|--------------|--------------------|
| 32 | 803 | 97.8 | 300 | 2 Q9S6R1 | Q9S6R1 escherichia |
| 33 | 798 | 97.2 | 300 | 2 Q6JKU8 | Q6JKU8 escherichia |
| 34 | 797 | 97.1 | 300 | 2 Q6JKU7 | Q6JKU7 escherichia |
| 35 | 714.5 | 87.0 | 243 | 2 Q08858 | Q08858 klebsiella |
| 36 | 501 | 61.0 | 129 | 2 Q9XBV8 | Q9XBV8 escherichia |
| 37 | 309.5 | 37.7 | 304 | 1 YDEQ_ECOLI | P77588 escherichia |
| 38 | 309.5 | 37.7 | 304 | 2 Q7AE58 | Q7AE58 escherichia |
| 39 | 309.5 | 37.7 | 304 | 2 Q8XAX2 | Q8XAX2 escherichia |
| 40 | 307.5 | 37.5 | 304 | 2 Q8CW27 | Q8CW27 escherichia |
| 41 | 195.5 | 23.8 | 245 | 2 Q83RF2 | Q83RF2 shigella fl |
| 42 | 128 | 15.6 | 299 | 2 Q53298 | Q53298 escherichia |
| 43 | 127 | 15.5 | 299 | 1 SFAH_ECOL6 | P13431 escherichia |
| 44 | 119 | 14.5 | 336 | 2 Q8CVB9 | Q8CVB9 escherichia |
| 45 | 117 | 14.3 | 299 | 2 Q46686 | Q46686 escherichia |

ALIGNMENTS

RESULT 1

Q9F5R9 PRELIMINARY; PRT; 166 AA.
AC Q9F5R9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia;
CX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Harris S.L., Spears P.A., Havell E.A., Hamrick T.S., Horton J.R.,
RA Orndorff P.E.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306537; AAG30927.1; -.
DR HSSP; P08191; 10QN.
DR InterPro; IPR008966; Adhes_bact.
SQ SEQUENCE 166 AA; 17826 MW; 2E64EEA90AE286DC CRC64;

Query Match 100.0%; Score 821; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 6.1e-67;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLNGWSVNAWSPACKTANGTAIPGGGSANVYNLAPVNVGNLVVDLS 60
Db 1 MKRVITLFAVLNGWSVNAWSPACKTANGTAIPGGGSANVYNLAPVNVGNLVVDLS 60

QY 61 TQIFCHNDYPTIIDYVTLQGSAYGGVLSNFSGTGVKSGSSYPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPTIIDYVTLQGSAYGGVLSNFSGTGVKSGSSYPPTTSETPRVYNSRT 120

QY 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 2

FIMH_ECOLI STANDARD; PRT; 300 AA.
AC P08191;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE FimH protein precursor.
GN Name=fimH; OrderedLocusNames=b4320;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562;

RN SEQUENCE FROM N.A.
 RP MEDLINE=88038337; PubMed=2890081;
 RX Klemm P., Christiansen G.;
 RA "Three fim genes required for the regulation of length and mediation
 RT of adhesion of *Escherichia coli* type 1 fimbriae.";
 RL Mol. Gen. Genet. 208:439-445(1987).
 RN [2] :
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=95334362; PubMed=7610040;
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blattner F.R.;
 RT "Analysis of the *Escherichia coli* genome VI: DNA sequence of the
 RT region from 92.8 through 100 minutes.";
 RL Nucleic Acids Res. 23:2105-2119(1995).
 RN [3] :
 RP CHARACTERIZATION.
 RX MEDLINE=90256291; PubMed=1971261;
 RA Krogfelt K.A., Bergmans H., Klemm P.;
 RT "Direct evidence that the FimH protein is the mannose-specific adhesin
 RT of *Escherichia coli* type 1 fimbriae.";
 RL Infect. Immun. 58:1995-1998(1990).
 CC -[1-] FUNCTION: Involved in regulation of length and mediation of
 CC adhesion of type 1 fimbriae (but not necessary for the production
 CC of fimbriae). Adhesin responsible for the binding to D-mannose. It
 CC is laterally positioned at intervals in the structure of the type
 CC 1 fimbriae. In order to integrate fimH in the fimbriae fimF and
 CC fimG are needed.
 CC -[1-] SIMILARITY: Belongs to the fimbrial protein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X05672; CAA29156.1; -;
 CC EMBL; U14003; AAB97216.1; -;
 CC EMBL; AE000502; AAC77276.1; -;
 CC F01; S56545; S56545;
 CC PDB; 1KIU; X-ray; B/D/F/H/J/L/N/P=22-300.
 CC PDB; 1KLF; X-ray; B/D/F/H/J/L/N/P=22-300.
 CC PDB; 1QUN; X-ray; B/D/F/H/J/L/N/P=22-300.
 CC DR ECHOBASE; EB0311; -;
 CC DR EcoGene; EG10315; fimH.
 CC DR InterPro; IPR008966; Adhes_bact.
 CC DR InterPro; IPR00259; Fimbrial.
 CC DR Pfam; PF00419; Fimbrial; 1.
 KW 3D-structure; Complete proteome; Fimbria; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 300
 FT FAIN FimH protein.
 FT CONFLICT 197 197
 FT CONFLICT 222 222 P -> R (in Ref. 1).
 FT STRAND 23 26 T -> H (in Ref. 1).
 FT TURN 27 28
 FT STRAND 31 32
 FT TURN 34 35
 FT STRAND 38 43
 FT STRAND 48 48
 FT TURN 51 52
 FT STRAND 55 58
 FT TURN 59 61
 FT HELIX 62 62
 FT TURN 63 66
 FT STRAND 70 72
 FT TURN 75 84
 FT STRAND 86 91
 FT TURN 92 98
 FT STRAND 99 100
 FT TURN 101 105
 FT STRAND 101 105

FT STRAND 114 116
 FT STRAND 122 123
 FT STRAND 126 132
 FT TURN 134 135
 FT STRAND 139 141
 FT TURN 143 144
 FT STRAND 146 156
 FT STRAND 163 171
 FT STRAND 175 176
 FT STRAND 183 183
 FT STRAND 190 191
 FT STRAND 200 201
 FT STRAND 205 207
 FT STRAND 212 219
 FT STRAND 222 222
 FT TURN 225 226
 FT STRAND 229 231
 FT STRAND 238 238
 FT STRAND 240 240
 FT TURN 241 241
 FT STRAND 242 248
 FT TURN 249 250
 FT STRAND 251 253
 FT TURN 255 256
 FT STRAND 259 264
 FT STRAND 269 270
 FT STRAND 273 279
 FT STRAND 285 285
 FT STRAND 288 288
 FT STRAND 291 292
 FT STRAND 294 300
 SQ SEQUENCE 300 AA; 31473 MW; 939204A51658747D CRC64;
 Query Match 100.0%; Score 821; DB 1; Length 300;
 Best Local Similarity 100.0%; Pred. No. 1.2e-66;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKRVITLFAVLMLGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
 DB 1 MKRVITLFAVLMLGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
 QY 61 TQIFCHNDYPTITDYVTLQSGAYGVLNFSGTGVKSGSSYPPTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPTITDYVTLQSGAYGVLNFSGTGVKSGSSYPPTTSETPRVYNSRT 120
 QY 121 DKFWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
 DB 121 DKFWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
 RESULT 3
 Q6JKV6 PRELIMINARY; PRT; 300 AA.
 ID Q6JKV6
 AC Q6JKV6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE FimH.
 GN Name=fimH;
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OC NCBI_TaxID=562;
 RX SEQUENCE FROM N.A.
 RX PubMed=15172698;
 RA Vandemaele F.J.;
 RT "Conservation of deduced amino acid sequence of FimH among *Escherichia*
 RT coli of bovine, porcine and avian disease origin.";
 RL Vet. Microbiol. 101:147-152(2004).
 RL EMBL; AY392512; AAR27559.1; -;
 DR InterPro; IPR008966; Adhes_bact.
 DR InterPro; IPR000259; Fimbrial.

```
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31457 MW; DC8CEBA50022C9BB CRC64;

Query Match      100.0%; Score 821; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-66;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSNWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSNWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120

QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

-RESULT 4
Q84DW2
ID Q84DW2 PRELIMINARY; PRT; 300 AA.
AC Q84DW2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DR 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE FimH
Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=562;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=22545382; PubMed=12657207;
RA Vandemaele F., Vandekerckhove D., Vereecken M., Derijcke J.,
Dho-Moulin M., Goddeeris B.M.;
"Sequence analysis demonstrates the conservation of fimH and
RT variability of fimA throughout avian pathogenic Escherichia coli
(RT (APEC).";
RL Vet. Res. 34:153-163 (2003).
RN [2]
SEQUENCE FROM N.A.
RP Vandemaele F.J., Goddeeris B.M.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF490857; AA084616.1; -
DR HSP: P08191; IQUIN
DR GO: GO:0002289; C:fimbria; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR008966; Adhes_bact.
DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31489 MW; 939219121PBF602D CRC64;

Query Match      100.0%; Score 821; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-66;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSNWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSNWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120

QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 5
```

```
Q9F5S0
ID Q9F5S0 PRELIMINARY; PRT; 300 AA.
AC Q9F5S0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE FimH
Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=562;
[1]
SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=21289098; PubMed=11395476;
RA Harris S.L., Spears P.A., Havell E.A., Hamrick T.S., Horton J.R.,
Orndorff P.E.;
"Characterization of Escherichia coli type 1 pilus mutants with
RT altered binding specificities.";
RL J. Bacteriol. 183:4099-4102 (2001).
DR EMBL: AF306536; AAG30926.1; -
DR HSP: P08191; IQUIN
DR GO: GO:0009289; C:fimbria; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR008966; Adhes_bact.
DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31413 MW; 922904051758745D CRC64;

Query Match      100.0%; Score 821; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-66;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSNWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSNWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120

QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 6
Q7BYN8
ID Q7BYN8 PRELIMINARY; PRT; 300 AA.
AC Q7BYN8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Minor fimbrial subunit, D-mannose specific adhesin.
Name=fimH; OrderedLocusNames=S4456;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
NCBI_TaxID=623;
[1]
SEQUENCE FROM N.A.
RP STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
Schwartz D.C., Blattner F.R.;
"Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL: AE016992; AAP19406.1; -
DR InterPro: IPR008966; Adhes_bact.
```

```

DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31445 MW; E0278E990BC74884 CRC64;

Query Match
Best Local Similarity 99.5%; Score 817; DB 2; Length 300;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSPACKTANGTAIPIGGSSANVYNLAPVNVGNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSPACKTANGTAIPIGGSSANVYNLAPVNVGNLVVDLS 60
QY 61 TQIFCHNDYPETITDVTTLQSGAYGGVLSNFGTVKYSYSSYPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDVTTLQSGAYGGVLSNFGTVKYSYSSYPPTTSETPRVYNSRT 120
QY -121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 7 .
Q83TY4 PRELIMINARY; PRT; 300 AA.
AC Q83TY4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22545392; PubMed=12657207;
RA Vandemaale F., Vanderkerchove D., Vereecken M., Derijcke J.,
RA Dho-Moulin M., Goddeeris B.M.;
RT "Sequence analysis demonstrates the conservation of fimH and
RT variability of fima throughout avian pathogenic Escherichia coli
RT (APEC).";
RL Vet. Res. 34:153-163(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Vandemaale F.J., Goddeeris B.M.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=15172698;
RA Vandemaale F.J., Hensen S.M., Goddeeris B.M.;
RT "Conservation of deduced amino acid sequence of FimH among Escherichia
RT coli of bovine, porcine and avian disease origin.";
RL Vet. Microbiol. 101:147-152(2004).
DR EMBL; AF490845; AAO84604.1; -
DR EMBL; AF490847; AAO84606.1; -
DR EMBL; AF490848; AAO84607.1; -
DR EMBL; AF490850; AAO84609.1; -
DR EMBL; AF490851; AAO84610.1; -
DR EMBL; AF490852; AAO84611.1; -
DR EMBL; AF490853; AAO84612.1; -
DR EMBL; AF490855; AAO84614.1; -
DR EMBL; AF490858; AAO84617.1; -
DR EMBL; AF490859; AAO84618.1; -
DR EMBL; AF490860; AAO84619.1; -
DR EMBL; AF490861; AAO84620.1; -
DR EMBL; AF490862; AAO84621.1; -
DR EMBL; AF490863; AAO84622.1; -
DR EMBL; AF490864; AAO84623.1; -
DR EMBL; AY392515; AAR27562.1; -
DR HSP; P08191; 1QUN.
DR GO; GO:000289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.

DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31445 MW; E0278E990BC74884 CRC64;

Query Match
Best Local Similarity 99.5%; Score 817; DB 2; Length 300;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSPACKTANGTAIPIGGSSANVYNLAPVNVGNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSPACKTANGTAIPIGGSSANVYNLAPVNVGNLVVDLS 60
QY 61 TQIFCHNDYPETITDVTTLQSGAYGGVLSNFGTVKYSYSSYPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDVTTLQSGAYGGVLSNFGTVKYSYSSYPPTTSETPRVYNSRT 120
QY -121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 7 .
Q83TY4 PRELIMINARY; PRT; 300 AA.
AC Q83TY4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22545392; PubMed=12657207;
RA Vandemaale F., Vanderkerchove D., Vereecken M., Derijcke J.,
RA Dho-Moulin M., Goddeeris B.M.;
RT "Sequence analysis demonstrates the conservation of fimH and
RT variability of fima throughout avian pathogenic Escherichia coli
RT (APEC).";
RL Vet. Res. 34:153-163(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Vandemaale F.J., Goddeeris B.M.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=15172698;
RA Vandemaale F.J., Hensen S.M., Goddeeris B.M.;
RT "Conservation of deduced amino acid sequence of FimH among Escherichia
RT coli of bovine, porcine and avian disease origin.";
RL Vet. Microbiol. 101:147-152(2004).
DR EMBL; AF490845; AAO84604.1; -
DR EMBL; AF490847; AAO84606.1; -
DR EMBL; AF490848; AAO84607.1; -
DR EMBL; AF490850; AAO84609.1; -
DR EMBL; AF490851; AAO84610.1; -
DR EMBL; AF490852; AAO84611.1; -
DR EMBL; AF490853; AAO84612.1; -
DR EMBL; AF490855; AAO84614.1; -
DR EMBL; AF490858; AAO84617.1; -
DR EMBL; AF490859; AAO84618.1; -
DR EMBL; AF490860; AAO84619.1; -
DR EMBL; AF490861; AAO84620.1; -
DR EMBL; AF490862; AAO84621.1; -
DR EMBL; AF490863; AAO84622.1; -
DR EMBL; AF490864; AAO84623.1; -
DR EMBL; AY392515; AAR27562.1; -
DR HSP; P08191; 1QUN.
DR GO; GO:000289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.

```

```

RESULT 9
Q83P73 ID Q83P73 PRELIMINARY; PRT; 300 AA.
AC Q83P73; 2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Minor fimbrial subunit, D-mannose specific adhesin.
GN Name=fim5; OrderedLocNames=SF4200;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP STRAIN=301 / Serotype 2a;
RC MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; AE015428; AAN45621.1; -.
DR HSP; P08191; 1QUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome.
SQ SEQUENCE 300 AA; 31445 MW; B0278E890BC74884 CRC64;

Query Match 99.5%; Score 817; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 2.8e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
DB 1 MKRAITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
DB 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 10
Q83P73 ID Q83P73 PRELIMINARY; PRT; 300 AA.
AC Q83P73; 2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Minor fimbrial subunit, D-mannose specific adhesin.
GN Name=fim5; OrderedLocNames=SF4200;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP STRAIN=301 / Serotype 2a;
RC MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; AE015428; AAN45621.1; -.
DR HSP; P08191; 1QUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome.
SQ SEQUENCE 300 AA; 31445 MW; B0278E890BC74884 CRC64;

Query Match 99.5%; Score 817; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 2.8e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
DB 1 MKRAITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
DB 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 11
Q83P73 ID Q83P73 PRELIMINARY; PRT; 300 AA.
AC Q83P73; 2003 (T-EMBLrel. 24, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP STRAIN=K12;
RC MEDLINE=20327582; PubMed=10869080;
RA Hamrick T.S., Harris S.L., Spears P.A., Havell E.A., Horton J.R.,
RA Russell P.W., Orndorff P.E.;
RT "Genetic characterization of Escherichia coli type 1 pilus adhesin
RT mutants and identification of a novel binding phenotype."
RL J. Bacteriol. 182:4012-4021(2000).
DR EMBL; AF154926; AAD44320.1; -.
DR HSP; P08191; 1QUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome.
SQ SEQUENCE 300 AA; 31499 MW; 3ADF97A85F8CE478 CRC64;

Query Match 99.3%; Score 815; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 4.2e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
DB 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158

```

```
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
RESULT 12
Q98497 ID Q98497 PRELIMINARY; PRT; 300 AA.
AC Q98497;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FimH.
OS Escherichia coli.
GN Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=20327582; PubMed=10869080;
RA Hamrick T.S., Harris S.L., Spears P.A., Havell E.A., Horton J.R.,
RA Russell P.W., Orndorff P.E.;
RT "Genetic characterization of Escherichia coli type 1 pilus adhesin
RT mutants and identification of a novel binding phenotype.";
RL J. Bacteriol. 182:4012-4021(2000).
DR EMBL; AF154925; AAD44319.1; -.
DR HSSP; P08191; 1QUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31517 MW; D29340A1575A2358 CRC64;
Query Match 99.3%; Score 815; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. NO. 4.2e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKRVITLFAVLGMGWSNWSACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
Db 1 MKRVITLFAVLGMGWSNWSACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
QY 61 TQIFCHNDYPETITDYVTLQSGAYSAGVLSNFSGTGKSGSSYPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQSGAYSAGVLSNFSGTGKSGSSYPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
RESULT 14
Q84DW4 ID Q84DW4 PRELIMINARY; PRT; 300 AA.
AC Q84DW4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE FimH.
OS Name=fimH;
GN Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=2255382; PubMed=12657207;
RA Vandenaële F., Vandekerchove D., Vereecken M., Derlijcke J.,
RA Dho-Moulin M., Goddeeris B.M.;
RT "Sequence analysis demonstrates the conservation of fimH and
RT variability of fimA throughout avian pathogenic Escherichia coli
RT (APEC).";
RL Vet. Res. 34:153-163(2003).
RN [2];
RP SEQUENCE FROM N.A.
RA Vandenaële F.J., Goddeeris B.M.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490849; AAC084608.1; -.
DR HSSP; P08191; 1QUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31472 MW; 8BDCCDB3F536AD64 CRC64;
Query Match 99.1%; Score 814; DB 2; Length 300;
Best Local Similarity 98.7%; Pred. NO. 5.2e-66;
Matches 156; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKRVITLFAVLGMGWSNWSACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
Db 1 MKRVITLFAVLGMGWSNWSACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
QY 61 TQIFCHNDYPETITDYVTLQSGAYSAGVLSNFSGTGKSGSSYPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQSGAYSAGVLSNFSGTGKSGSSYPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
RESULT 15
Q985Y2 ID Q985Y2 PRELIMINARY; PRT; 300 AA.
AC Q985Y2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE FIMH SUBUNIT=MANNOSSE-sensitive type 1 fimbrial adhesin.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=94131954; PubMed=7905476;
RA Sokurenko E.V., Courtney H.S., Ohman D.E., Klemm P., Hasty D.L.;
RL J. Bacteriol. 176:748-755(1994).
DR HSSP; P08191; 1QUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
```

Q9F5S1
ID Q9F5S1 PRELIMINARY; PRT; 300 AA.
AC Q9F5S1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE FimH.
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=21289098; PubMed=11395476;
RA Harris S.L., Spears P.A., Havell E.A., Hamrick T.S., Horton J.R.,
RA Orndorff P.E.;
RT "Characterization of Escherichia coli type 1 pilus mutants with
RT altered binding specificities.";
RL J. Bacteriol. 183:4099-4102(2001).
DR EMBL; AF306535; AAC30925.1; -
DR HSP; P08191; IQUN
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes_bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31457 MW; 9F7B0A35FF51F938 CRC64;

Query Match 99.1%; Score 814; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 5.2e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MKRVITLFAVLLMGVSNVWSPACKTANGTAIPIGGSSAVVYNLAPVNVGNLVVDLS 60
DB 1 MKRVITLFAVLLMGVSNVWSPACKTANGTAIPIGGSSAVVYNLAPVNVGNLVVDLS 60

OY 61 TQIFCHNDYPTITDYVTLQGSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRT 120
DB 61 TQIFCHNDYPTITDYVTLQGSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRT 120

OY 121 DKWPVALYITPVSSAGGVAIKAGSLIAVLIRQTNKY 158
DB 121 DKWPVALYITPVSSAGGVAIKAGSLIAVLIRQTNKY 158

Search completed: November 27, 2004, 16:10:47
Job time : 204 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2004, 16:06:28 ; Search time 39 Seconds
(without alignments)
389.801 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158
Perfect score: 821
Sequence: 1 MKRVTLFAVLLMGWSNAN.....VAIKAGSLIAVLLRQTNVY 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 821 | 100.0 | 300 | S56545 | fimbrial protein f |
| 2 | 811 | 98.8 | 300 | G91017 | hypothetical prote |
| 3 | 811 | 98.8 | 300 | B86130 | hypothetical prote |
| 4 | 714.5 | 87.0 | 302 | A32801 | fimbrial adhesin p |
| 5 | 309.5 | 37.7 | 304 | A64304 | probable fimbrial |
| 6 | 309.5 | 37.7 | 304 | F85725 | probable adhesin, |
| 7 | 309.5 | 37.7 | 304 | C90892 | probable adhesin (|
| 8 | 128 | 15.6 | 299 | C49233 | S fimbrial adhesin |
| 9 | 127 | 15.5 | 291 | S15927 | sfaH protein precu |
| 10 | 117 | 14.3 | 299 | I76900 | F1652 minor fimbri |
| 11 | 86.5 | 10.5 | 682 | JQ0420 | beta-1,3-glucanase |
| 12 | 85 | 10.4 | 317 | T39869 | probable lysophosp |
| 13 | 82.5 | 10.0 | 354 | A82693 | uroporphyrinogen d |
| 14 | 82.5 | 10.0 | 408 | T22801 | hypothetical prote |
| 15 | 80.5 | 9.8 | 473 | A59200 | acid phosphatase (|
| 16 | 80.5 | 9.8 | 482 | C84330 | pantothenate perme |
| 17 | 80.5 | 9.8 | 558 | B99494 | dihydroxy-acid deh |
| 18 | 80 | 9.7 | 1114 | JH0284 | 125K surface anti |
| 19 | 79.5 | 9.7 | 432 | T71683 | NADH2 dehydrogenas |
| 20 | 79.5 | 9.7 | 653 | A60961 | leishmanolysin (EC |
| 21 | 79.5 | 9.7 | 824 | A50767 | S-receptor kinase |
| 22 | 78.5 | 9.6 | 343 | H95879 | probable sugar ABC |
| 23 | 78.5 | 9.6 | 451 | C78932 | hypothetical prote |
| 24 | 78 | 9.5 | 184 | C86320 | hypothetical prote |
| 25 | 78 | 9.5 | 282 | AB1359 | biogenesis of thyl |
| 26 | 78 | 9.5 | 471 | G64665 | hypothetical prote |
| 27 | 77.5 | 9.4 | 412 | S72579 | hypothetical prote |
| 28 | 77.5 | 9.4 | 1337 | B64993 | hypothetical prote |
| 29 | 77.5 | 9.4 | 1534 | A85862 | hypothetical prote |

| | | | | | | |
|----|------|-----|------|---|--------|--------------------|
| 30 | 77.5 | 9.4 | 1534 | 2 | G91017 | probable membrane |
| 31 | 77 | 9.4 | 309 | 2 | AD0221 | probable sugar ABC |
| 32 | 77 | 9.4 | 471 | 2 | B71849 | probable outer mem |
| 33 | 77 | 9.4 | 630 | 2 | S37663 | leptomeric polypro |
| 34 | 77 | 9.4 | 1154 | 1 | VGIHIB | E2 glycoprotein pr |
| 35 | 76 | 9.3 | 246 | 2 | C70527 | hypothetical prote |
| 36 | 75.5 | 9.2 | 378 | 2 | S00842 | leukosialin precu |
| 37 | 75.5 | 9.2 | 383 | 2 | G84248 | iron-binding prote |
| 38 | 75.5 | 9.2 | 485 | 2 | T51094 | acid phosphatase (|
| 39 | 75 | 9.1 | 276 | 2 | G86130 | hypothetical prote |
| 40 | 75 | 9.1 | 276 | 2 | D91289 | hypothetical prote |
| 41 | 75 | 9.1 | 849 | 2 | D82510 | chitinase VCA0027 |
| 42 | 75 | 9.1 | 1041 | 2 | S55862 | probable membrane |
| 43 | 74.5 | 9.1 | 341 | 2 | E96019 | probable sugar upt |
| 44 | 74.5 | 9.1 | 352 | 2 | A95098 | hypothetical prote |
| 45 | 74.5 | 9.1 | 396 | 2 | AE2137 | hypothetical prote |

ALIGNMENTS

RESULT 1

S56545
fimbrial protein fimb precursor, type 1 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S56545; B65246; S09563; A36967
R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A>Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.
A:Reference number: S56314; MUID:95334362; PMID:7610040
A:Accession: S56545
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-300 <BLAT>
A:Cross-references: UNIPROT:P08191; EMBL:U14003; NID:G1263172; PIDN:AAA97216.1; PID:9537
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, August 1994
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B65246
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-300 <BLAT>
A:Cross-references: GB:AE000502; GB:U00096; NID:G2367374; PIDN:AAC77276.1; PID:G1790775;
A:Experimental source: strain K-12, substrain MG1655
R:Klemm, P.; Christiansen, G.
Mol. Gen. Genet. 208, 439-445, 1987
A>Title: Three fim genes required for the regulation of length and mediation of adhesion
A:Reference number: S07321; MUID:88038337; PMID:2890081
A:Accession: S09563
A:Molecule type: DNA
A:Residues: 1-196, 'R', 198-221, 'H', 223-300 <KLE>
A:Cross-references: EMBL:X05672; NID:G41463; PIDN:CAA29156.1; PID:G41466
R:Sokurenko, E.V.; Courtney, H.S.; Ohman, D.E.; Klemm, P.; Hasty, D.L.
J. Bacteriol. 176, 748-755, 1994
A>Title: FimH family of type 1 fimbrial adhesins: functional heterogeneity due to minor
A:Reference number: A36967; MUID:94131954; PMID:7905476
A:Accession: A36967
A>Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-78, 'R', 80-221, 'H', 223-300 <SOK>
A>Note: sequence extracted from NCBI backbone (NCBIP:143314)
C:Genetics:
A:Gene: fimH
C:Function:
A:Description: involved in longitudinal regulation and mannose-specific adhesion
A>Note: not necessary for the production of fimbriae
A>Note: controls length and number of fimbriae
C:Superfamily: fimbrial protein fimbH
C:Keywords: fimbria

F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-300/Product: fimbrial protein fimH, type 1 #status

```
Query Match      100.0%; Score 821; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.3e-68;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVVNLPVVNVGNLVVDLS 60

1 MKRVITLFAVLLMGWSVNAWSEACKTANGTAIPIGGGSANVYNLAPVVNVGONLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTTLQGSAYGGVLSNFSGTVKYSGSSYPFFTSETPRVVYNSRT 120

db 61 TOIFCHNDYPETITDXYVTLORGSAYGGVLSNFSGTVKYSGSSYPFPPTSETPRVVNSRT 120

QY 121 DKF:WPVALYLTTPVSSAGGVAIKAGSLIAVLILROTNYY 158

Db 12 F: DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILPOTNNY 158

RESULT 2

G91288 .
Gyodetrical protein ECS279 [imported] - Escherichia coli (strain O157:H7, substrate R11
O Gyodetrical protein ECS279 [imported] - Escherichia coli (strain O157:H7, substrate R11

C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

[illegible]

gasawara, N.; Makino, K.; Omisaka, M.; Kurokawa, K.; Ishii, K.; Tokoyama, K.; Han, C.-G. DNA Res. 8, 11-22, 2001

A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic analysis of the *hly* and *eae* genes
 JNA Res. 8, 11-22, 2001
 3.Reference numbers: 300520, MIMD, 01155031, DMD, 11269702

A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: G91288

A; Status: preliminary
A'; Molecule type: DNA

A:Residues: 1-300 <HAY>
A:Cross-references: UNIPROT:O8XBA6; GB:EA000007; PDB:BA38702.1; PID:q13664757; CSPDB:G

A; Experimental source: strain O157:H7, substrain RMD 0509952
C; Genetics:
A; cross-references: UNIPROT:Q9AB46; GBT:SA000007; PIDN:BAE38702.1; FID:G13364757; GSPDB:1

Gene: EC55279
Superfamily: fimbrial protein fimH
Query Match 98.8% Score 811 DB 2 Length 300

| | | | |
|-----------------------|--------------|--------------------|-----------|
| Best Local Similarity | 98.7%; | Pred: No. 4.4e-67; | |
| Matches 156: | Conservative | 0: Mismatches | 2: Indels |

QV 1 MKBYTTE EAYTJ MCW SIANWSEACVTANCTA TDTCCCANWZANWY A DYTATYCONY VETD S CA

1 MEBVITTI EAST I MCVSINIAWCEACVANCE T BTCCCCANTENTY A DARTCONT FRY C

61 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 14

61

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

DD

RESULT 3

hypotheetical protein fimH [imported] - Escherichia coli (strain O157:H7 substrate EDL933)
B86130

C.Species: Escherichia coli
 C.Date: 16-Feb-2001 #sequence revision 16-Feb-2001
 C.Date: 09-Jul-2004 #text change 09-Jul-2004

C:\Access: B86130
C:\Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

Ry:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamensis, K.; Apodaca, M.

A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A;Residues: 1-304 <HAY>
A;Cross-references: UNIPROT:Q8XAY2; GB:BA000007; PIDN:BAB35630.1; PID:gl13361573; GSPDB:B
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: ECs2107
C;Superfamily: fimbrial protein fimH

Query Match 37.7%; Score 309.5; DB 2; Length 304;
Best Local Similarity 44.1%; Pred.No. 5.7e-21;
Matches 63; Conservative 29; Mismatches 48; Indels 3; Gaps 2;

QY 7 LFAVLLMGMSVANVSFACKTANGTAIPICGGSANVVNLAPVVNVGNQLVVDLSTQTIFCHN-DYP 66
Db 10 LFGIYLLMAGKYFAFSNCVNDGSS--IGAGITSVVNLDPVIQPCONLVLDLSQHISCM 67
QY 67 NDYPERI-TDYVTIQGSAYGVLNFSGTSVKYSGSSYPPTTSETPRVVNSRTDKWP 125
Db 68 NDYGGWYTDHINLVQSAPAGSLQSKYGLNNVTVFPPLTTNTNVLDIGDKTPMLP 127
QY 126 VALYLTPVSSAGGVAIKAGSLIA 148
Db 128 LKLYIIPVGAGGVVIKAGEVIA 150

RESULT 8
C49233
S fimbrial adhesin minor subunit SfaH - Escherichia coli
C;Species: Escherichia coli
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: C49233
R;Hacker, J.; Kestler, H.; Hoschutsky, H.; Jann, K.; Lottspeich, F.; Korhonen, T.K.
Infect. Immun. 61, 544-550, 1993
A;Title: Cloning and characterization of the S fimbrial adhesin II complex of an Escherichia coli
A;Reference number: A49233; MUID:93138776; PMID:8093693
A;Contents: 018:K1
A;Accession: C49233
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-299 <HAC>
A;Cross-references: UNIPROT:Q53298; GB:S53210; NID:G264033; PIDN:AAB25047.1; PID:g264036
A;Note: sequence extracted from NCBI backbone (NCBIN:123181, NCBIP:123186)
C;Superfamily: fimbrial protein fimH

Query Match 15.8%; Score 128; DB 2; Length 299;
Best Local Similarity 25.8%; Pred.No. 0.00027;
Matches 39; Conservative 24; Mismatches 78; Indels 10; Gaps 4;

QY 13 MGMSVANVSFACKTANGTAIPICGGSANVVNLAPVVNVGNQL-VVDLSTQTIFCHN-DYP 70
Db 1 MAYSQPSFALLCRN-NQTGOVFNSGDTSRVNVSPVVOYDKSIVLDDLSQLSCQNEST 59
QY 71 ETITDVYTLQRSAYGGLVS-----NFTSGTVKYSGSSYPPTTSETPRVVNSRTDKP 123
Db 60 QONYDYLIKILKGFGFPALDTKTGYRLDTSRPTGYARQLPFLQDLQVTEAFYQYGVWKP 119
QY 124 WPVALYITPVSSAGGVAIKAGSLIAVILLRQ 154
Db 120 FPAKLYLPAPGVFGKVINGNDLLATLYNK 150

RESULT 9
S15927
sfah protein precursor - Escherichia coli
C;Species: Escherichia coli
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C;Accession: S15927; S06195
R;Schmoll, T.; Hoschutsky, H.; Morschhauser, J.; Lottspeich, F.; Jann, K.; Hacker, J.
Mol. Microbiol. 3, 1735-1744, 1989
A;Title: Analysis of genes coding for the sialic acid-binding adhesin and two other minor
A;Reference number: S15925; MUID:90158121; PMID:2576095
A;Accession: S15927
A;Molecule type: DNA
A;Residues: 1-291 <SCH>

A;Cross-references: UNIPROT:PI3431; EMBL:X16664; NID:g42955; PIDN:CAA34654.1; PID:g42962
C;Genetics:
A;Gene: sfah
C;Superfamily: fimbrial protein fimH
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-291/Product: sfah protein #status predicted <MAT>

Query Match 15.5%; Score 127; DB 2; Length 291;
Best Local Similarity 25.8%; Pred. No. 0.00032;
Matches 39; Conservative 24; Mismatches 78; Indels 10; Gaps 4;

QY 13 MGWSNVNMFACKTANGTAIPIGGSANVYVNLAPVNVGNQL-VVDLSTQIFCHN-DYP 70
Db 1 MAYSQSFALLCRN-NQTQFNSGDTSPRVNVPVVEYDKSISVLDLSQLVSCQNEB 59
QY 71 ETITDYVTLQGSAYGGVLS-----NFSGTVKYSGSSYPPTTSETPRVYNSRTDKP 123
Db 60 GQNDYLYKILKSGFSPALDNTYGRDLPTSRPTGVARQLPQLDQVTEAFYQYGVNKP 119
QY 124 WPAVLYLTPVSSAGGVAIKAGSLIAVLILRQ 154
Db 120 FPAKWLYPEPGVFGKVINNGDGLATLYVVK 150

RESULT 10
I76900
F1852 minor fimbrial chain H - Escherichia coli
C;Species: Escherichia coli
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C;Accession: I76900; I80334
R;Harel, J.; Jacques, M.; Fairbrother, J.M.; Bosse, M.; Forget, C.
Microbiology 141, 221-228, 1995
A;Title: Cloning of determinants encoding F185(2) fimbriae from porcine septicaemic Esch
A;Reference number: I59111; MUID:95202083; PMID:7894716
A;Accession: I76900
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-299 <RES>
A;Cross-references: UNIPROT:O46686; EMBL:U09804; NID:g967126; PIDN:AAA74946.1; PID:g9671
R;van Die, I.; Kramer, C.; Hacker, J.; Bergmans, H.; Jongen, W.; Hoekstra, W.
Res. Microbiol. 142, 653-658, 1991
A;Title: Nucleotide sequence of the genes coding for minor fimbrial subunits of the F1C
A;Reference number: I59446; MUID:92073661; PMID:1683712
A;Accession: I80334
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-299 <RE2>
A;Cross-references: GB:S68237; NID:g239708; PIDN:AAS20439.1; PID:g239711
C;Genetics:
A;Gene: foch
C;Superfamily: fimbrial protein fimH

Query Match 14.3%; Score 117; DB 2; Length 299;
Best Local Similarity 25.8%; Pred. No. 0.0028;
Matches 39; Conservative 27; Mismatches 75; Indels 10; Gaps 6;

QY 13 MGWSNVNMFACKTANGTAIPIGGSANVYVNLAPVNVGNQL-VVDLSTQIFCHN-DYP 70
Db 1 MAYSQSFALLCRN-NQTQFNSGDTSPRVNVPVVEYDKSISVLDLSQLVSCQNEB 59
QY 71 ETITDYVTLQGSAYGGVLS-----NFSGTVKYSGSSYPPTTSETPRVYNSRTDKP 123
Db 60 GQNDYLYKILKSGFSPALDNTYGRDLPTSRPTGVARQLPQLDQVTEAFYQYGVNKP 119
QY 124 WPAVLYLTPVSSAGGVAIKAGSLIAVLILRQ 154
Db 120 FPAKWLYPEPGVFGKVINNGDGLATLYVVK 150

RESULT 11
J00420
beta-1,3-glucanase A1 precursor - Bacillus circulans
C;Species: Bacillus circulans

A;Cross-references: UNIPROT:P23903; GB:M34503; NID:g142972; PIDN:AAA22474.1; PID:g142973
A;Residues: 1-682 <VOS>
A;Cross-references: UNIPROT: strain WL-12
A;Experimental source: strain WL-12
C;Comment: This enzyme, together with chitinase, is crucial for hydrolyzing yeast and fur
C;Genetics:
A;Gene: glcA
F;1-38/Domain: signal sequence #status predicted <SIG>
F;39-682/Product: beta-1,3-glucanase A1 #status predicted <MAT>

Query Match 10.5%; Score 86.5; DB 2; Length 682;
Best Local Similarity 22.2%; Pred. No. 4.7;
Matches 47; Conservative 30; Mismatches 56; Indels 79; Gaps 9;

QY 1 MKRVITLFAVLLMGWSNVNMFACKTANGTAI-----PI----- 34
Db 12 MKKVLGLFLVWMLASVGLVPTSKVQAAGTIVTSMYFSPADGPVLSKSGVKASYGVN 71
QY 35 ---GGGSA---NYYVNLAPVNVGNQLV-VD-----LSTQIFCH-----NDYPETITDY 76
Db 72 PKFNGSATWVDYSDGVNVKGVNWDIDQAGVIYQNVNWHWSDDGFGNHWTLST 131
QY 77 VTLQGRSAYGGV-----LSNFSGTVKYSGSSYPPTT 109
Db 132 TEIQLYSKANGVLEQLVFNQINKTTITAMNPTQGPQTASFTG-----GAGFTYPTEN 186
QY 110 ETRPVYNSRTDPPWPAVALYLPVSSAGGVAI 141
Db 187 NDSAVTYEAVD---DLKVVYKPVNSSSWIDI 215

RESULT 12
T39869
Probable lysophospholipase (EC 3.1.1.5) SPBC1E8.05 [similarity] - fission yeast (Schizos
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39869; T42012
R;Volckaert, G.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21887
A;Accession: T39869
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-317 <VOL>
A;Cross-references: UNIPROT:O42970; EMBL:AL021746; NID:g2842520; PIDN:CAA16838.1; PID:g28
A;Experimental source: strain 972h; cosmid c1E8
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T42012
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 6-123, 'K' 125-317, 'IVVS' <VOS>
A;Cross-references: EMBL:D89103; NID:g1749413; PIDN:BAAL3766.1; PID:g1749414
A;Experimental source: strain PR745
C;Genetics:
A;Gene: SPDB:SPBC1E8.05
A;Map position: 2
C;Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase
C;Keywords: carboxylic ester hydrolase

Query Match 10.4%; Score 85; DB 2; Length 317;
Best Local Similarity 26.0%; Pred. No. 2.6;
Matches 38; Conservative 24; Mismatches 46; Indels 38; Gaps 7;

```
QY 9 AVLLMGVNVNWSFACKTANGTAIP--GGGSANVYNLAPVNVVGNQL-----VVDLS 60
Db 27 AVLQEGGVNTVSNLSTVTLTYRGNSA-----LTPETIASDIDNTGYLWNIA 81
QY 61 TQIFCHNDY-----PETIDVVTIQRGS-----AYGGVLSNFS-----CTVK 97
Db 82 TYTAAADYLLGLSFDGGETYSQYFTIQAAGTCTISTSSLSYSGTSSSTIASRSMIGT-- 139
QY 98 YSGSSYFPPTTSEPRVNVNSRTDKP 123
Db 140 RTSSSYFTTSSSTPSSSSSSSSSSP 165

RESULT 13
A82693
uoroporphyrinogen decarboxylase XF1332 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82693
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82615; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-354 <SLM>
A:Cross-references: UNIPROT:Q9PDP7; GB:A5003966; GB:A5003849; NID:99106327; PIDN:AAF8414
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.B.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XPI332
C:Superfamily: uroporphyrinogen decarboxylase

Query Match 10.0%; Score 82.5; DB 2; Length 354;
Best Local Similarity 23.2%; Pred. No. 5;
Matches 44; Conservative 24; Mismatches 57; Indels 65; Gaps 10;

QY 9 AVLLMGVNVNWSFACKTANGTAIPGGGSANVYNL-----APVNVGNLWVDLSQ 62
Db 136 SVFLIGSGSPWTLAC-----YMEGGSKSYAHKAMAFNAP-----DVLHQ 178

QY 63 IFCHNDYPETIDVVTIQRGS-----AYGGVLSNFSGVTKYSGSSYPPT----- 107
Db 179 LL--NTVNAVSYLAQAAGAAQALQVFDTWGGILS-----PMYRTSLPYLTRAREL 232

QY 108 -----TSEPRVNVN-----SRTDK-----PWPVALYITPVSSAGGVAIKAGSLI 147
Db 233 ERGTGERTPLVFGKNGBYIAELAMSGTEAVGVVDVTIELEDAARRSNGVALQ--GNLD 291

QY 148 AVLILRQTN 157
Db 292 PATLYGTFFNN 301

RESULT 14
T22801
hypotheical protein F56H6.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
```

```
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22801
R:Kershaw, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19617
A:Accession: T22801
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-408 <WIL>
A:Cross-references: UNIPROT:O45586; EMBL:Z81553; PIDN:CAB04497.1; GSPDB:GN00019; CESP:F56
A:Experimental source: clone F56H6
C:Genetics:
A:Gene: CESP:F56H6.8
A:Map position: 1
A:Introns: 67/3; 159/1

Query Match 10.0%; Score 82.5; DB 2; Length 408;
Best Local Similarity 28.8%; Pred. No. 6;
Matches 32; Conservative 16; Mismatches 44; Indels 19; Gaps 5;

QY 24 CKTANGTAIPIG--GGSANVYNLAPVNVVGNQL--VVDLSQIFCHNDYP-----E 71
Db 287 CSMTSTTAIPPEIPSCNVCVGMAGPVINSRLSGSVNCTYQLATLGPYKIALKFTFFD 346

QY 72 TITDVTIQRG-SAYGGVLSNFSCTVKYSGSSYPPTTSETPRVNVNSRTD 121
Db 347 TGNDIYTIYDGPSTESPILG-----RYSNNMYEFTTSGSTMVVTFKSD 391
```

```
RESULT 15
A59200
acid phosphatase (EC 3.1.3.2) purple 1, precursor [validated] - sweet potato
N:Alternate names: purple acid phosphatase (PAP)
C:Species: Ipomoea batatas (sweet potato)
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: A59200; C59200
R:Schenk, G.; Ge, Y.; Carrington, L.E.; Wynne, C.J.; Searle, I.R.; Carroll, B.J.; Hamilt
Arch. Biochem. Biophys. 370, 183-189, 1999
A:Title: Binuclear metal centers in plant purple acid phosphatases: Fe-Mn in sweet potat
A:Reference number: Z25293; MUID:99441212; PMID:10510276
A:Accession: A59200
A:Molecule type: mRNA
A:Residues: 1-473 <KUR>
A:Cross-references: UNIPROT:Q9SE00; GB:AF200825; NID:G6635440; PIDN:AAF19821.1; PID:G663
A:Accession: C59200
A:Molecule type: protein
A:Residues: 39-61 <KU2>
C:Comment: Metal content is iron and manganese, with small amounts of iron and copper.
C:Complex: homodimer
C:Function:
A:Description: catalyzes the hydrolysis of phosphate monoesters [validated, MUID:9944121
A:Note: active on a variety of activated and unactivated phosphate ester and anhydride s
C:Superfamily: kidney bean purple acid phosphatase; phosphatase core homology
F:39-473/Product: purple acid phosphatase #status experimental <MAT>
F:166-240/Domain: phosphoesterase core homology <PEC>
F:118,180,311,433/Binding site: carboxylate (Asn) (covalent) #status predicted
F:172,201,204,362/Binding site: iron (Asp, Asp, Tyr, His) #status predicted
F:201,238,323,360/Binding site: manganese (Asp, Asp, His, His) #status predicted
F:239,333/Active site: His #status predicted
F:382/Disulfide bonds: interchain #status predicted
```

```
Query Match 9.8%; Score 80.5; DB 1; Length 473;
Best Local Similarity 37.9%; Pred. No. 11;
Matches 22; Conservative 12; Mismatches 19; Indels 5; Gaps 2;
```

```
QY 76 YVTIQRGSAYGGVLSNFSCTVKYSGSSYPPTTSETPRVNVNSRTDWPVALYITPV 133
Db 274 WYAIKRASAHITVLSYSGFVKYSPQYKWF--TSELEKV---NRSETPWILVLAHPL 326
```

Search completed: November 27, 2004, 16:07:18
Job time : 43 secs

us-10-015-085-4_copy_1_158.rpr

Mon Nov 29 14:12:21 2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2004, 16:07:26 ; Search time 141 Seconds

(without alignments)
397.459 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158

Perfect score: 821

Sequence: 1 MKRVITLFAVLLMGWSVNAW.....VAIKAGSLIAVLRLQNTNNY 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

-Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

-Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 821 | 100.0 | 300 | 9 | US-09-912-020-367 |
| 2 | 821 | 100.0 | 300 | 14 | US-10-288-978-2 |
| 3 | 821 | 100.0 | 300 | 14 | US-10-015-085-4 |
| 4 | 821 | 100.0 | 300 | 17 | US-10-681-381B-1 |
| 5 | 809 | 98.5 | 300 | 15 | US-10-607-834-14 |
| 6 | 712 | 86.7 | 268 | 10 | US-09-747-802-1 |
| 7 | 712 | 86.7 | 268 | 16 | US-10-789-619-1 |
| 8 | 712 | 86.7 | 279 | 9 | US-09-900-575-34 |
| 9 | 712 | 86.7 | 279 | 9 | US-09-900-575-44 |
| 10 | 709 | 86.4 | 279 | 9 | US-09-900-575-42 |
| 11 | 708 | 86.2 | 279 | 9 | US-09-900-575-27 |
| 12 | 708 | 86.2 | 279 | 9 | US-09-900-575-36 |
| 13 | 708 | 86.2 | 279 | 9 | US-09-900-575-39 |

| | | | | | | |
|----|-----|------|-----|----|------------------|-------------------|
| 14 | 708 | 86.2 | 279 | 9 | US-09-900-575-55 | Sequence 55, Appl |
| 15 | 706 | 86.0 | 279 | 9 | US-09-900-575-23 | Sequence 23, Appl |
| 16 | 705 | 85.9 | 279 | 9 | US-09-900-575-29 | Sequence 29, Appl |
| 17 | 704 | 85.7 | 279 | 9 | US-09-900-575-32 | Sequence 32, Appl |
| 18 | 704 | 85.7 | 279 | 9 | US-09-900-575-37 | Sequence 37, Appl |
| 19 | 703 | 85.6 | 280 | 9 | US-09-900-575-30 | Sequence 30, Appl |
| 20 | 702 | 85.5 | 279 | 9 | US-09-900-575-43 | Sequence 43, Appl |
| 21 | 700 | 85.3 | 279 | 9 | US-09-900-575-31 | Sequence 31, Appl |
| 22 | 700 | 85.3 | 279 | 9 | US-09-900-575-38 | Sequence 38, Appl |
| 23 | 700 | 85.3 | 279 | 9 | US-09-900-575-40 | Sequence 40, Appl |
| 24 | 700 | 85.3 | 279 | 15 | US-10-607-834-15 | Sequence 15, Appl |
| 25 | 697 | 84.9 | 279 | 9 | US-09-900-575-25 | Sequence 25, Appl |
| 26 | 697 | 84.9 | 279 | 9 | US-09-900-575-26 | Sequence 26, Appl |
| 27 | 697 | 84.9 | 279 | 9 | US-09-900-575-33 | Sequence 33, Appl |
| 28 | 697 | 84.9 | 279 | 9 | US-09-900-575-45 | Sequence 45, Appl |
| 29 | 694 | 84.5 | 279 | 9 | US-09-900-575-24 | Sequence 24, Appl |
| 30 | 694 | 84.5 | 279 | 9 | US-09-900-575-28 | Sequence 28, Appl |
| 31 | 694 | 84.5 | 279 | 9 | US-09-900-575-35 | Sequence 35, Appl |
| 32 | 690 | 84.0 | 279 | 9 | US-09-900-575-41 | Sequence 41, Appl |
| 33 | 130 | 15.8 | 46 | 10 | US-09-747-802-74 | Sequence 74, Appl |
| 34 | 130 | 15.8 | 46 | 16 | US-10-789-619-74 | Sequence 74, Appl |
| 35 | 128 | 15.6 | 46 | 10 | US-09-747-802-75 | Sequence 75, Appl |
| 36 | 128 | 15.6 | 46 | 16 | US-10-789-619-75 | Sequence 75, Appl |
| 37 | 123 | 15.0 | 27 | 10 | US-09-747-802-2 | Sequence 2, Appl |
| 38 | 123 | 15.0 | 27 | 16 | US-10-789-619-2 | Sequence 2, Appl |
| 39 | 122 | 14.9 | 25 | 10 | US-09-747-802-3 | Sequence 3, Appl |
| 40 | 122 | 14.9 | 25 | 16 | US-10-789-619-3 | Sequence 3, Appl |
| 41 | 118 | 14.4 | 46 | 10 | US-09-747-802-77 | Sequence 77, Appl |
| 42 | 118 | 14.4 | 46 | 16 | US-10-789-619-77 | Sequence 77, Appl |
| 43 | 111 | 13.5 | 46 | 10 | US-09-747-802-76 | Sequence 76, Appl |
| 44 | 111 | 13.5 | 46 | 16 | US-10-789-619-76 | Sequence 76, Appl |
| 45 | 102 | 12.4 | 25 | 10 | US-09-747-802-4 | Sequence 4, Appl |

ALIGNMENTS

RESULT 1

US-09-912-020-367
; Sequence 367, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Chlsen, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; FILE REFERENCE: ESCHERICHIA COLI
; CURRENT APPLICATION NUMBER: US/09/912,020
; CURRENT FILING DATE: 2001-07-23
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/117,405
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 367
; LENGTH: 300
; TYPE: PRT
; ORGANISM: E. Coli
US-09-912-020-367

Query Match 100.0%; Score 821; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWVFACKTANGTAIP:GGGSANVYVNLAPVVGQNLVVDLS 60
|||||

Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 2

US-10-288-978-2
; Sequence 2, Application US/10288978
; Publication No. US20030099665A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Hultgren, Scott J.
; APPLICANT: Pinkner, Jerome S.
; TITLE OF INVENTION: Chapterone and Adhesin Proteins; Vaccines, Diagnostics
; FILE OF INVENTION: and Methods for Treating Infections
; FILE REFERENCE: 469201-362
; CURRENT APPLICATION NUMBER: US/10/288,978
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: US/09/298,494
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 60/082,824
; PRIOR FILING DATE: 1998-04-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-288-978-2

Query Match 100.0%; Score 821; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 3

US-10-015-085-4
; Sequence 4, Application US/10015085
; Publication No. US20030199071A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon R.
; APPLICANT: Hultgren, Scott J.
; APPLICANT: Hung, Chia-Suei
; APPLICANT: Bouckaert, Julie
; TITLE OF INVENTION: Mutant Proteins, High Potency Inhibitory Antibodies, and FimCH
; FILE OF INVENTION: Crystall Structure
; FILE REFERENCE: 10271-037
; CURRENT APPLICATION NUMBER: US/10/015,085
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 300
; TYPE: PRT
; ORGANISM: E. coli

US-10-015-085-4

Query Match 100.0%; Score 821; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 4

US-10-681-381B-1
; Sequence 1, Application US/10681381B
; Publication No. US20040224400A1
; GENERAL INFORMATION:
; APPLICANT: Schembri, Mark A
; APPLICANT: Klemm, Per
; TITLE OF INVENTION: Novel multifunctional adhesin proteins and their display in
; FILE OF INVENTION: microbial cells
; FILE REFERENCE: 54259.000004
; CURRENT APPLICATION NUMBER: US/10/681,381B
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US 60/083,794
; PRIOR FILING DATE: 1998-05-01
; PRIOR APPLICATION NUMBER: DK 1998 00598
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: US 09/301,704
; PRIOR FILING DATE: 1999-04-29
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-681-381B-1

Query Match 100.0%; Score 821; DB 17; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 5

US-10-607-834-14
; Sequence 14, Application US/10607834
; Publication No. US20040067544A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, PC
; APPLICANT: Vogel, Viola
; TITLE OF INVENTION: Use of Adhesion Molecules as Bond Stress-Enhanced Nanoscale Bindir
; FILE OF INVENTION: Switches
; FILE REFERENCE: 91-02
; CURRENT APPLICATION NUMBER: US/10/607,834


```
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US 60/392,467
; PRIOR FILING DATE: 2002-06-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 31
; SEQ ID NO 14
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-607-834-14

Query Match      98.5%; Score 809; DB 15; Length 300;
Best Local Similarity 98.1%; Pred. No. 3.8e-75;
Matches 155; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60

QY 61 TQIFCHNDYPETITDVTYIQRGSAYGVLNFSGTVKYSGSYPPPTTSETPRVYNSRT 120
DB 61 TQIFCHNDYPETITDVTYIQRGSAYGVLNFSGTVKYSGSYPPPTTSETPRVYNSRT 120

QY 121 DKRPVVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158
DB 121 DKRPVVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 6
US-09-747-802-1
; Sequence 1, Application US/09747802
; Publication No. US20030027979A1
; GENERAL INFORMATION:
; APPLICANT: WANG, CHANG YI
; TITLE OF INVENTION: SYNTHETIC PEPTIDE COMPOSITION AS IMMUNOGENS FOR
; FILE REFERENCE: 1151-4165
; CURRENT APPLICATION NUMBER: US/09/747,802
; CURRENT FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-747-802-1

Query Match      86.7%; Score 712; DB 10; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.6e-65;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPETITDVTYIQR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPETITDVTYIQR 60

QY 82 GSAYGVLNFSGTVKYSGSYPPPTTSETPRVYNSRTDKPWPVALYITPVSSAGGVAI 141
DB 61 GSAYGVLNFSGTVKYSGSYPPPTTSETPRVYNSRTDKPWPVALYITPVSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 7
US-10-789-619-1
; Sequence 1, Application US/10789619
; Publication No. US20040141993A1
; GENERAL INFORMATION:
; APPLICANT: WANG, CHANG YI
; TITLE OF INVENTION: SYNTHETIC PEPTIDE COMPOSITION AS IMMUNOGENS FOR
; FILE REFERENCE: 1151-4165
```

```
; CURRENT APPLICATION NUMBER: US/10/789,619
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-789-619-1

Query Match      86.7%; Score 712; DB 16; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.6e-55;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPETITDVTYIQR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPETITDVTYIQR 60

QY 82 GSAYGVLNFSGTVKYSGSYPPPTTSETPRVYNSRTDKPWPVALYITPVSSAGGVAI 141
DB 61 GSAYGVLNFSGTVKYSGSYPPPTTSETPRVYNSRTDKPWPVALYITPVSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 8
US-09-900-575-34
; Sequence 34, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-34

Query Match      86.7%; Score 712; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 3.8e-65;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPETITDVTYIQR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPETITDVTYIQR 60

QY 82 GSAYGVLNFSGTVKYSGSYPPPTTSETPRVYNSRTDKPWPVALYITPVSSAGGVAI 141
DB 61 GSAYGVLNFSGTVKYSGSYPPPTTSETPRVYNSRTDKPWPVALYITPVSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 9
US-09-900-575-44
; Sequence 44, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
```

APPLICANT: Revel, Andrew
APPLICANT: Auguste, Christine
APPLICANT: Burlein, Jeanne
TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
FILE REFERENCE: 469201-549
CURRENT APPLICATION NUMBER: US/09/900,575
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/60/216,750
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.0
SEQ ID NO 44
LENGTH: 279
TYPE: PRT
ORGANISM: E. coli
US-09-900-575-44

Query Match 86.7%; Score 712; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 3.8e-65;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 FACTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSQIFCHNDYPETITDYVTLQR 81
DB 1 FACTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSQIFCHNDYPETITDYVTLQR 60
QY 82 GSAYGGVLSNFGSTGVKSGSSYPFPTTSETPRVYNSTRDKWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTGVKSGSSYPFPTTSETPRVYNSTRDKWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 10
US-09-900-575-42
Sequence 42, Application US/09900575
Patent No. US20020150587A1
GENERAL INFORMATION:
APPLICANT: Langermann, Solomon
APPLICANT: Revel, Andrew
APPLICANT: Auguste, Christine
APPLICANT: Burlein, Jeanne
TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
FILE REFERENCE: 469201-549
CURRENT APPLICATION NUMBER: US/09/900,575
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/60/216,750
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.0
SEQ ID NO 42
LENGTH: 279
TYPE: PRT
ORGANISM: E. coli
US-09-900-575-42

Query Match 86.4%; Score 709; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.8e-65;
Matches 136; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 22 FACTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSQIFCHNDYPETITDYVTLQR 81
DB 1 FACTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSQIFCHNDYPETITDYVTLQR 60
QY 82 GSAYGGVLSNFGSTGVKSGSSYPFPTTSETPRVYNSTRDKWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTGVKSGSSYPFPTTSETPRVYNSTRDKWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 11
US-09-900-575-27
Sequence 27, Application US/09900575
Patent No. US20020150587A1
GENERAL INFORMATION:
APPLICANT: Langermann, Solomon
APPLICANT: Revel, Andrew
APPLICANT: Auguste, Christine
APPLICANT: Burlein, Jeanne
TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
FILE REFERENCE: 469201-549
CURRENT APPLICATION NUMBER: US/09/900,575
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/60/216,750
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 279
TYPE: PRT
ORGANISM: E. coli
US-09-900-575-27

Query Match 86.2%; Score 708; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 9.9e-65;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 22 FACTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSQIFCHNDYPETITDYVTLQR 81
DB 1 FACTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSQIFCHNDYPETITDYVTLQR 60
QY 82 GSAYGGVLSNFGSTGVKSGSSYPFPTTSETPRVYNSTRDKWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTGVKSGSSYPFPTTSETPRVYNSTRDKWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 12
US-09-900-575-36
Sequence 36, Application US/09900575
Patent No. US20020150587A1
GENERAL INFORMATION:
APPLICANT: Langermann, Solomon
APPLICANT: Revel, Andrew
APPLICANT: Auguste, Christine
APPLICANT: Burlein, Jeanne
TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
FILE REFERENCE: 469201-549
CURRENT APPLICATION NUMBER: US/09/900,575
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/60/216,750
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.0
SEQ ID NO 36
LENGTH: 279
TYPE: PRT
ORGANISM: E. coli
US-09-900-575-36

Query Match 86.2%; Score 708; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 9.9e-65;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 22 FACTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSQIFCHNDYPETITDYVTLQR 81
DB 1 FACTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSQIFCHNDYPETITDYVTLQR 60
QY 82 GSAYGGVLSNFGSTGVKSGSSYPFPTTSETPRVYNSTRDKWPVALYLTVPSSAGGVAI 141

Db 61 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVYNSRTDKPWPVALYLTTPVSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 13
US-09-900-575-39
; Sequence 39, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-39

Query Match 86.2%; Score 708; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 9.9e-65;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYNLAPVNVGNLWVLDLSTQIFCHNDYPETITDYVTLQR 81
Db 1 FACKTANGTAIPIGGGSANVYNLAPVNVGNLWVLDLSTQIFCHNDYPETITDYVTLQR 60
QY 82 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVYNSRTDKPWPVALYLTTPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVYNSRTDKPWPVALYLTTPVSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 14
US-09-900-575-55
; Sequence 55, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 55
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence of FimH proteins for SEQ ID NO: 23 to 45
US-09-900-575-55

Query Match 86.2%; Score 708; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 9.9e-65;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 22 FACKTANGTAIPIGGGSANVYNLAPVNVGNLWVLDLSTQIFCHNDYPETITDYVTLQR 81
Db 1 FACKTANGTAIPIGGGSANVYNLAPVNVGNLWVLDLSTQIFCHNDYPETITDYVTLQR 60
QY 82 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVYNSRTDKPWPVALYLTTPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVYNSRTDKPWPVALYLTTPVSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 15
US-09-900-575-23
; Sequence 23, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-23

Query Match 86.0%; Score 706; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 1.6e-64;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 22 FACKTANGTAIPIGGGSANVYNLAPVNVGNLWVLDLSTQIFCHNDYPETITDYVTLQR 81
Db 1 FACKTANGTAIPIGGGSANVYNLAPVNVGNLWVLDLSTQIFCHNDYPETITDYVTLQR 60
QY 82 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVYNSRTDKPWPVALYLTTPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVYNSRTDKPWPVALYLTTPVSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

Search completed: November 27, 2004, 16:24:26
Job time : 142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2004, 16:06:28 ; Search time 39 Seconds
(without alignments)
268.673 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158
Perfect score: 821
Sequence: 1 MKRVITLFAVLLMGWSVNAW.....VAIKAGSLIATVILRQTNNY 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

.Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

.Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/protdata/1/iaa/SA COMB.pep.*
2: /cgn2_6/protdata/1/iaa/SB COMB.pep.*
3: /cgn2_6/protdata/1/iaa/SA COMB.pep.*
4: /cgn2_6/protdata/1/iaa/SB COMB.pep.*
5: /cgn2_6/protdata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/protdata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 821 | 100.0 | 300 | 4 | US-09-298-494-2 |
| 2 | 821 | 100.0 | 300 | 4 | US-09-492-709A-367 |
| 3 | 714.5 | 87.0 | 314 | 4 | US-09-489-039A-11890 |
| 4 | 712 | 86.7 | 288 | 4 | US-09-747-802-1 |
| 5 | 712 | 86.7 | 279 | 4 | US-09-900-575-34 |
| 6 | 712 | 86.7 | 279 | 4 | US-09-900-575-44 |
| 7 | 709 | 86.4 | 279 | 4 | US-09-900-575-42 |
| 8 | 708 | 86.2 | 279 | 4 | US-09-900-575-27 |
| 9 | 708 | 86.2 | 279 | 4 | US-09-900-575-36 |
| 10 | 708 | 86.2 | 279 | 4 | US-09-900-575-39 |
| 11 | 708 | 86.2 | 279 | 4 | US-09-900-575-55 |
| 12 | 706 | 86.0 | 279 | 4 | US-09-900-575-23 |
| 13 | 705 | 85.9 | 279 | 4 | US-09-900-575-29 |
| 14 | 704 | 85.7 | 279 | 4 | US-09-900-575-32 |
| 15 | 704 | 85.7 | 279 | 4 | US-09-900-575-37 |
| 16 | 703 | 85.6 | 280 | 4 | US-09-900-575-30 |
| 17 | 702 | 85.5 | 279 | 4 | US-09-900-575-43 |
| 18 | 700 | 85.3 | 279 | 4 | US-09-900-575-31 |
| 19 | 700 | 85.3 | 279 | 4 | US-09-900-575-38 |
| 20 | 700 | 85.3 | 279 | 4 | US-09-900-575-40 |
| 21 | 697 | 84.9 | 279 | 4 | US-09-900-575-25 |
| 22 | 697 | 84.9 | 279 | 4 | US-09-900-575-26 |
| 23 | 697 | 84.9 | 279 | 4 | US-09-900-575-33 |
| 24 | 697 | 84.9 | 279 | 4 | US-09-900-575-45 |
| 25 | 694 | 84.5 | 279 | 4 | US-09-900-575-28 |
| 26 | 694 | 84.5 | 279 | 4 | US-09-900-575-24 |
| 27 | 694 | 84.5 | 279 | 4 | US-09-900-575-35 |

28 690 84.0 279 4 US-09-900-575-41 Sequence 41, Appl
29 246 30.0 317 4 US-09-543-681A-5763 Sequence 5763, Ap
30 232.5 28.3 303 4 US-09-543-681A-7946 Sequence 7946, Ap
31 130 15.8 46 4 US-09-747-802-74 Sequence 74, Appl
32 128 15.6 46 4 US-09-747-802-75 Sequence 75, Appl
33 123 15.0 27 4 US-09-747-802-2 Sequence 2, Appl
34 122 14.9 25 4 US-09-747-802-3 Sequence 3, Appl
35 118 14.4 46 4 US-09-747-802-77 Patent No. 5202113
36 115 14.0 28 6 5202113-3
37 111 13.5 46 4 US-09-747-802-76 Sequence 76, Appl
38 102 12.4 25 4 US-09-747-802-4 Sequence 4, Appl
39 91.5 11.1 21 4 US-09-747-802-88 Sequence 88, Appl
40 82.5 10.0 338 4 US-09-328-352-5051 Sequence 5051, Ap
41 82 10.0 562 4 US-09-543-681A-4371 Sequence 4371, Ap
42 77.5 9.4 422 4 US-09-270-767-61739 Sequence 61739, A
43 75.5 9.2 499 2 US-09-032-315-3 Sequence 3, Appl
44 75.5 9.2 499 2 US-08-993-318A-3 Sequence 3, Appl
45 75.5 9.2 499 3 US-09-399-886-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-298-494-2
; Sequence 2, Application US/09298494
; Patent No. 6500434

GENERAL INFORMATION:

; APPLICANT: Langermann, Solomon
; APPLICANT: Hultgren, Scott J.
; APPLICANT: Pinkner, Jerome S.
; TITLE OF INVENTION: Chaperone and Adhesin Proteins; Vaccines, Diagnostics
; TITLE OF INVENTION: and Methods for Treating Infections
; FILE REFERENCE: 469201-362
; CURRENT APPLICATION NUMBER: US/09/298,494
; CURRENT FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: US 60/082,824
; EARLIER FILING DATE: 1998-04-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-298-494-2

Query Match 100.0%; Score 821; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.1e-81;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKRVITLFAVLLMGWSVNAWFACTANGTAIPIGGGSANYVNLAPVNVYQNLVVDLS 60
DB 1 MKRVITLFAVLLMGWSVNAWFACTANGTAIPIGGGSANYVNLAPVNVYQNLVVDLS 60
QY 61 TQICHNDYPTITDYTLQKSAYGVLNFSFSTVKYSGSYFPFTTSETPRVYVNSRT 120
DB 61 TQICHNDYPTITDYTLQKSAYGVLNFSFSTVKYSGSYFPFTTSETPRVYVNSRT 120
QY 121 DKPFPVALYLPVSSAGGVAIKAGSLIATVILRQTNNY 158
DB 121 DKPFPVALYLPVSSAGGVAIKAGSLIATVILRQTNNY 158

RESULT 2

US-09-492-709A-367
; Sequence 367, Application US/09492709A
; Patent No. 6720139

; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.

```
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION: ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001A
; CURRENT APPLICATION NUMBER: US/09/492,709A
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 367
; LENGTH: 300
; TYPE: PRT
; ORGANISM: E. Coli
; US-09-492-709A-367

Query Match      100.0%; Score 821; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.1e-81;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MKRVITLPAVLGMGNSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
Db      1 MKRVITLPAVLGMGNSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60

Qy      61 TQIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTGKSGSSYPFPPTTSETPRVYNSRT 120
Db      61 TQIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTGKSGSSYPFPPTTSETPRVYNSRT 120

Qy      121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILROTNYY 158
Db      121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILROTNYY 158

RESULT 3
US-09-489-039A-11890
; Sequence 11890, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11890
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-11890

Query Match      87.0%; Score 714.5; DB 4; Length 314;
Best Local Similarity 84.3%; Pred. No. 1.7e-69;
Matches 134; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

Qy      1 MKRVITLPAVLGMGNSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDL 59
Db      14 MKKIIPLFTLLGLGNSNAWSPACKTATGTAIPIGGGSANVYNLTPAVNVGQNLVVDL 73

Qy      60 TQIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTGKSGSSYPFPPTTSETPRVYNSR 119
Db      74 STQIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTGKSGSSYPFPPTTSETARVIYDSR 133

Qy      120 TDKPWPVALYLPVSSAGGVAIKAGSLIAVLILROTNYY 158
Db      134 TDKPWPVALYLPVSTAGGVAITAGSLIAVLILROTNYY 172

RESULT 4
US-09-747-802-1
; Sequence 1, Application US/09747802
```

```
; Patent No. 6780969
; GENERAL INFORMATION:
; APPLICANT: WANG, CHANG YI
; TITLE OF INVENTION: SYNTHETIC PEPTIDE COMPOSITION AS IMMUNOGENS FOR
; TITLE OF INVENTION: PREVENTION OF URINARY TRACT INFECTION
; FILE REFERENCE: 1151-4165
; CURRENT APPLICATION NUMBER: US/09/747,802
; CURRENT FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-747-802-1
```

```
Query Match      86.7%; Score 712; DB 4; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.5e-69;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      22 FACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS TQIFCHNDYPETITDVTTLOR 81
Db      1 FACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS TQIFCHNDYPETITDVTTLOR 60

Qy      82 GSAYGGVLSNFSGTGKSGSSYPFPPTTSETPRVYNSRTDKPWPVALYLPVSSAGGVAI 141
Db      61 GSAYGGVLSNFSGTGKSGSSYPFPPTTSETPRVYNSRTDKPWPVALYLPVSSAGGVAI 120

Qy      142 KAGSLIAVLILROTNYY 158
Db      121 KAGSLIAVLILROTNYY 137
```

```
RESULT 5
US-09-900-575-34
; Sequence 34, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
; US-09-900-575-34
```

```
Query Match      86.7%; Score 712; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.7e-69;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      22 FACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS TQIFCHNDYPETITDVTTLOR 81
Db      1 FACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS TQIFCHNDYPETITDVTTLOR 60

Qy      82 GSAYGGVLSNFSGTGKSGSSYPFPPTTSETPRVYNSRTDKPWPVALYLPVSSAGGVAI 141
Db      61 GSAYGGVLSNFSGTGKSGSSYPFPPTTSETPRVYNSRTDKPWPVALYLPVSSAGGVAI 120

Qy      142 KAGSLIAVLILROTNYY 158
Db      121 KAGSLIAVLILROTNYY 137
```

```
RESULT 6
US-09-900-575-44
; Sequence 44, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlain, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 44
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-44

Query Match      86.7%; Score 712; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.7e-69;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLOR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLOR 60

QY 82 GSAYGGVLSNFGSTVKYSGSSYPPTTSETPRVYVNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTVKYSGSSYPPTTSETPRVYVNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNYY 158
DB 121 KAGSLIAVLILRQTNYY 137

RESULT 7
US-09-900-575-42
; Sequence 42, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlain, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 42
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-42

Query Match      86.4%; Score 709; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 5.7e-69;
Matches 136; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLOR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLOR 60

QY 82 GSAYGGVLSNFGSTVKYSGSSYPPTTSETPRVYVNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTVKYSGSSYPPTTSETPRVYVNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNYY 158
DB 121 KAGSLIAVLILRQTNYY 137

RESULT 8
US-09-900-575-27
; Sequence 27, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlain, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 27
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-27

Query Match      86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLOR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLOR 60

QY 82 GSAYGGVLSNFGSTVKYSGSSYPPTTSETPRVYVNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTVKYSGSSYPPTTSETPRVYVNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNYY 158
DB 121 KAGSLIAVLILRQTNYY 137

RESULT 9
US-09-900-575-36
; Sequence 36, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlain, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 36
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-36

Query Match      86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 60
QY 82 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 120
QY 142 KAGSLIAVLILROTNYY 158
Db 121 KAGSLIAVLILROTNYY 137

RESULT 10
US-09-900-575-39
; Sequence 39, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 39
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-39

Query Match 86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 60
QY 82 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 120
QY 142 KAGSLIAVLILROTNYY 158
Db 121 KAGSLIAVLILROTNYY 137

RESULT 11
US-09-900-575-55
; Sequence 55, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 55
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-55

Query Match 86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 60
QY 82 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 120
QY 142 KAGSLIAVLILROTNYY 158
Db 121 KAGSLIAVLILROTNYY 137

RESULT 12
US-09-900-575-23
; Sequence 23, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 23
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-23

Query Match 86.0%; Score 706; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 1.2e-68;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 60
QY 82 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 120
QY 142 KAGSLIAVLILROTNYY 158
Db 121 KAGSLIAVLILROTNYY 137

RESULT 13
US-09-900-575-29
; Sequence 29, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 29
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-29

Query Match 86.0%; Score 706; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 1.2e-68;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVYTLOR 60
QY 82 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPVSSAGGVAI 120
QY 142 KAGSLIAVLILROTNYY 158
Db 121 KAGSLIAVLILROTNYY 137
```



```

; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-29

```

| | | | | | | | | |
|-----------------------|--------------|---|------------|----|--------|--------|------|----|
| Query Match | 85.9%; | Score | 705; | DB | 4; | Length | 279; | |
| Best Local Similarity | 98.5%; | Pred. No. | 1.5e-69; | | | | | |
| Matches 135; | Conservative | 1; | Mismatches | 1; | Indels | 0; | Gaps | 0; |
| Qy | 22 | FAKKTANGTAIPIGGGSANVYNLAPVNVYGQNLVLDLSTQIFCHNDYDETITDYYVTQ | 81 | | | | | |
| Db | 1 | FAKKTANGTAIPIGGGSANVYNLAPVNVYGQNLVLDLSTQIFCHNDYDETITDYYVTQ | 60 | | | | | |
| Qy | 82 | GAAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPFPVALYILTVPSSAGGVAI | 141 | | | | | |
| Db | 61 | GAAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPFPVALYILTVPSSAGGLVI | 120 | | | | | |
| Qy | 142 | KAGSLIAVLIIROQTNNY | 158 | | | | | |
| Db | 121 | KAGSLIAVLIIROQTNNY | 137 | | | | | |

```

RESULT 14
US-09-900-575-32
; Sequence 32, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-32

```

| | Query Match | 85.7%; | Score 704; | DB 4; | Length 279; |
|----|-----------------------|--|------------------|-----------|-------------|
| | Best Local Similarity | 98.5%; | Pred. No. 2e-68; | | |
| | Matches 135; | Conservative 0; | Mismatches 2; | Indels 0; | Gaps 0; |
| Qy | 22 | FACKTANGTAIPIGGSSANVYNLAPVYVQNLVWDLSTQIECHNDYPETITDVTYLQR | 81 | | |
| Db | 1 | FACKTANGTAIPIGGSSANVYNLAPVYVQNLVWDLSTQIECHNDYPETITDVTYLQR | 60 | | |
| Qy | 82 | GSAYGGVLSNFSGTWKYSGSSYPPTTSETPRVYVNSRTDKPWPVALYILTPVSSAGGVAI | 141 | | |
| Db | 61 | GSAYGGVLSNFSGTWKYSGSSYPPTTSETPRVYVNSRTDKPWPVALYILTPVSSAGGVVI | 120 | | |
| Qy | 142 | KAGSLI AVLILROTNNY | 158 | | |
| Db | 121 | KAGSLI AVLILROTNNY | 137 | | |

RESULT 15
US-09-900-575-37
; Sequence 37. Application US/09900575

```

; PATENT NO. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langemann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burslein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
; US-09-900-575-37

```

| | Query Match | 85.7%; | Score 704; | DB 4; | Length 279; |
|----|-----------------------|---------------------------------------|---------------------------|-------------|-------------|
| | Best Local Similarity | 98.5%; | Pred. No. 2e-68; | | |
| | Matches 135; | Conservative 1; | Mismatches 1; | Indels 0; | Gaps 0; |
| Qy | 22 | FACKTANGTAIPGGGSANVYNLAPVWVNGQNLVVDLS | QIIFCHNDYPTETI | DYVTLQ | 81 |
| | | | | | |
| Db | 1 | FACKTANGTAIPGGGSANVYNLAPVWVNGQNLVVDLS | QIIFCHNDYPTETI | DYVTLQ | 60 |
| | | | | | |
| Qy | 82 | GSAYGGVLSNFSGTVKYSGSSYPFP | TTSETPRVYVNSRTDKPWPVALYLT | TPVSSAGGVAL | 141 |
| | | | | | |
| Db | 61 | GSAYGGVLSNFSGTVKYSGSSYPFP | TTSETPRVYVNSRTDKPWPVALYLT | TPVSSAGGVAL | 120 |
| | | | | | |
| Qy | 142 | KAGSLIAVLILROTNNY | 158 | | |
| | | | | | |
| Db | 121 | KAGSLIAVLILROTNNY | 137 | | |
| | | | | | |

Search completed: November 27, 2004, 16:14:07
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein -2protein search, using sw model

Run on: November 27, 2004, 16:06:28 ; Search time 153 Seconds
(without alignments)
370.452 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158

Perfect score: 821

Sequence: 1 MKRVITLFAVLWGWSVNAW.....VAIKAGSLIAVLILRTNNY 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesecp23Sep04:*

- 1: Genesecp1980s:*
- 2: Genesecp1990s:*
- 3: Genesecp2000s:*
- 4: Genesecp2001s:*
- 5: Genesecp2002s:*
- 6: Genesecp2003as:*
- 7: Genesecp2003bs:*
- 8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 821 | 100.0 | 300 | 2 | AAR76745 FimH prot |
| 2 | 821 | 100.0 | 300 | 3 | AAY59456 E. coli p |
| 3 | 821 | 100.0 | 300 | 3 | AAB16009 E. coli p |
| 4 | 821 | 100.0 | 300 | 5 | AAR77488 Escherich |
| 5 | 821 | 100.0 | 300 | 6 | ABU08998 E. coli a |
| 6 | 821 | 100.0 | 300 | 7 | ADC24381 E. coli F |
| 7 | 821 | 100.0 | 300 | 7 | ADG17420 Escherich |
| 8 | 817 | 99.5 | 300 | 2 | AAR76774 FimH prot |
| 9 | 817 | 99.5 | 300 | 2 | AAR76771 FimH prot |
| 10 | 817 | 99.5 | 300 | 2 | AAR76763 FimH prot |
| 11 | 816 | 99.4 | 300 | 2 | AAR76773 FimH prot |
| 12 | 815 | 99.3 | 300 | 7 | ADG17467 Escherich |
| 13 | 815 | 99.3 | 300 | 2 | AAR76768 FimH prot |
| 14 | 811 | 98.8 | 300 | 2 | AAR76772 FimH prot |
| 15 | 809 | 98.5 | 300 | 2 | AAR76770 FimH prot |
| 16 | 809 | 98.5 | 300 | 2 | AAR76775 FimH prot |
| 17 | 809 | 98.5 | 300 | 2 | AAR76764 FimH prot |
| 18 | 809 | 98.5 | 300 | 8 | ADH56426 Escherich |
| 19 | 808 | 98.4 | 300 | 2 | AAR76769 FimH prot |
| 20 | 802 | 97.7 | 300 | 2 | AAR76776 FimH prot |
| 21 | 801 | 97.6 | 300 | 2 | AAR76765 FimH prot |
| 22 | 797 | 97.1 | 300 | 2 | AAR76766 FimH prot |
| 23 | 787 | 95.9 | 296 | 2 | AAR76767 FimH prot |
| 24 | 714.5 | 87.0 | 314 | 7 | AB065373 Klebsiell |
| 25 | 712 | 86.7 | 268 | 5 | AB068154 Escherich |

| | | | | | |
|----|-----|------|-----|---|--------------------|
| 26 | 712 | 86.7 | 279 | 4 | AAV72515 Escherich |
| 27 | 712 | 86.7 | 279 | 4 | AB72839 Bacterial |
| 28 | 712 | 86.7 | 279 | 4 | AAB47074 Adhesin p |
| 29 | 712 | 86.7 | 279 | 5 | AAE18424 Escherich |
| 30 | 712 | 86.7 | 279 | 5 | AAE18433 Escherich |
| 31 | 712 | 86.7 | 296 | 4 | AAB47073 Immunogen |
| 32 | 712 | 86.7 | 304 | 4 | AAB47072 Immunogen |
| 33 | 709 | 86.4 | 279 | 5 | AAE18435 Escherich |
| 34 | 708 | 86.2 | 279 | 5 | AAE18429 Escherich |
| 35 | 708 | 86.2 | 279 | 5 | AAE18417 Escherich |
| 36 | 708 | 86.2 | 279 | 5 | AAE18436 Escherich |
| 37 | 708 | 86.2 | 279 | 5 | AAE18426 Escherich |
| 38 | 706 | 86.0 | 279 | 5 | AAE18413 Escherich |
| 39 | 705 | 85.9 | 279 | 5 | AAE18419 Escherich |
| 40 | 704 | 85.7 | 279 | 5 | AAE18427 Escherich |
| 41 | 704 | 85.7 | 279 | 5 | AAE18422 Escherich |
| 42 | 703 | 85.6 | 280 | 5 | AAE18420 Escherich |
| 43 | 702 | 85.5 | 279 | 5 | AAE18432 Escherich |
| 44 | 700 | 85.3 | 279 | 5 | AAE18430 Escherich |
| 45 | 700 | 85.3 | 279 | 5 | AAE18428 Escherich |

ALIGNMENTS

RESULT 1

| | | | | | |
|----------|---|-------------------------|--|--|--|
| AAR76745 | AAR76745 standard; protein; 300 AA. | | | | |
| XX | AC | AAR76745; | | | |
| XX | AC | AAR76745; | | | |
| DT | 16-OCT-2003 | (revised) | | | |
| DT | 13-MAR-1996 | (first entry) | | | |
| XX | | | | | |
| DE | FimH protein derived from E. coli K12 strain PC31. | | | | |
| XX | | | | | |
| XW | FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue; | | | | |
| XW | FimA; FimF; FimG; receptor binding site. | | | | |
| XX | | | | | |
| OS | Escherichia coli K12; strain PC31. | | | | |
| XX | | | | | |
| PH | Key | Location/Qualifiers | | | |
| FT | Peptide | 1..21 | | | |
| FT | | /note= "Signal peptide" | | | |
| FT | Protein | 22..300 | | | |
| FT | | /note= "Mature FimH" | | | |
| XX | | | | | |
| PN | WO9520657-A1. | | | | |
| XX | | | | | |
| PD | 03-AUG-1995. | | | | |
| XX | | | | | |
| PF | 27-JAN-1995; 95WO-DK000042. | | | | |
| XX | | | | | |
| PR | 27-JAN-1994; 94US-00187166. | | | | |
| XX | | | | | |
| PA | (GXBI-) GX BIOSYSTEMS AS. | | | | |
| XX | | | | | |
| PI | Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S; | | | | |
| XX | | | | | |
| DR | WPI; 1995-275442/36. | | | | |
| XX | | | | | |

Receptor specific bacterial adhesins - useful for targeting active compounds and microbial cells to locations of receptors.

Example 1; Page 88-89; 152pp; English.

This sequence represents the FimH protein from E. coli K12 strain PC31. FimH is located at the tip of the type 1 fimbriae and also intercalated at intervals in the fimbrial organelle. Most forms of the FimH adhesin target, and bind to, oligosaccharide structures containing terminally located alpha-D-mannoside residues. FimH contains 4 cysteine residues assumed to direct folding of the molecule into distinct functional domains. For comparison FimA and the minor components FimF and FimG only

CC have 2 cysteine residues. The localisation of the cysteine residues in
 CC FimH points to a tandem arrangement of two ancestral genes. Similar amino
 CC acids can be found in similar positions in the two halves of the FimH
 CC protein. The "midway" point is located roughly around residue 150 in the
 CC mature protein. The two halves or domains of FimH have evolved
 CC differently with the N-terminal section becoming the domain harbouring
 CC the receptor binding site, whereas the C-terminal sector became the
 CC domain of the molecule required for integration into the fimbrial
 CC organelle. This sequence and those given in AAR76763-76 may be used in
 CC the production of a variant FimH adhesin which may be useful for
 CC targeting active compounds and microbial cells to locations comprising
 CC selected receptors to which the adhesins bind. (Updated on 16-OCT-2003 to
 CC standardise OS field)

XX Sequence 300 AA;
 SQ
 Query Match 100.0%; Score 821; DB 2; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2e-79;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKRVITLFAVLLMGWSVNAWSACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
 DB 1 MKRVITLFAVLLMGWSVNAWSACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
 QY 61 TQIFCHNDYPETITDVTTLQGSAYGGVLSNFSGTVKYSYSGSYPPPTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPETITDVTTLQGSAYGGVLSNFSGTVKYSYSGSYPPPTTSETPRVYNSRT 120
 QY 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158
 DB 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 2
 AAY59456
 ID AAY59456 standard; peptide; 300 AA.
 XX
 AC AAY59456;
 XX
 DT 29-MAR-2000 (first entry)
 XX
 DE E. coli PC31 FimH protein.
 XX
 KW Multifunctional adhesin protein; organic receptor; bioremediation;
 KW biosorption; organic pollutant; herbicide; pesticide; toxic compound;
 KW recycling; metal isolation; metal binding domain.
 XX
 OS Escherichia coli.
 XX
 PN WO9957276-A1.
 XX
 PD 11-NOV-1999.
 XX
 PF 21-APR-1999; 99MO-DK000223.
 XX
 PR 30-APR-1998; 98DK-00000598.
 PR 01-MAY-1998; 98US-0083794P.
 XX
 XX (GYRE-) GYRE LTD.
 PA
 XX Schembri MA, Klemm P;
 PI
 XX WPI; 2000-072233/06.
 XX
 XX Novel recombinant cells useful for bioremediation and recycling
 XX processes.
 XX
 PS Claim 5; Page 5; 60pp; English.
 XX
 XX This sequence is the E. coli FimH protein, which is an adhesin protein.
 XX The invention relates to a recombinant cell expressing a multifunctional
 CC adhesin (MA) protein on its surface. The MA protein has at least one
 CC binding domain (BD1) capable of binding to an organic receptor, and at

CC least one binding domain (BD2) not naturally present in the adhesin, and
 CC can bind to a compound to which the naturally occurring adhesin protein
 CC does not substantially bind. Cells of the invention are used as
 CC bioremediation or biosorption means to separate undesired compounds such
 CC as organic pollutants including herbicides and pesticides, or toxic
 CC compounds such as heavy metals from the environment, or for isolating
 CC precious compounds such as metals for recycling purposes

XX Sequence 300 AA;
 SQ

Query Match 100.0%; Score 821; DB 3; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2e-79;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
 DB 1 MKRVITLFAVLLMGWSVNAWSACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
 QY 61 TQIFCHNDYPETITDVTTLQGSAYGGVLSNFSGTVKYSYSGSYPPPTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPETITDVTTLQGSAYGGVLSNFSGTVKYSYSGSYPPPTTSETPRVYNSRT 120
 QY 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158
 DB 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 3
 AAB16009
 ID AAB16009 standard; protein; 300 AA.
 XX
 AC AAB16009;
 XX
 DT 05-OCT-2000 (first entry)
 XX
 DE E. coli proliferation associated protein sequence SEQ ID NO:367.
 XX
 KW Escherichia coli; E. coli; proliferation; inhibition; screening;
 KW antimicrobial; bacterial growth; antisense therapy; antibacterial.
 XX
 OS Escherichia coli.
 XX
 EN WO2000044906-A2.
 XX
 PD 03-AUG-2000.
 XX
 PF 27-JAN-2000; 2000MO-US002200.
 XX
 PR 27-JAN-1999; 99US-0117405P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Zyskind J, Ohlssen KL, Trawick J, Forsyth RA, Froelich JM;
 PI Carr GJ, Yamamoto RT, Xu HH;
 XX
 XX WPI; 2000-514822/46.
 DR N-PSDB; AAB66015.
 XX
 XX Novel polynucleotides and polypeptides associated with microorganism
 PT proliferation, used to identify inhibitors of bacterial growth and
 PT proliferation, for use in antisense therapy.

XX Claim 11; Page 274-275; 316pp; English.
 XX
 XX AAB65809 to AAB65889 and AAB66058 to AAB66138 represent nucleotide
 CC sequences derived from Escherichia coli which inhibit E. coli
 CC proliferation. AAB65890 to AAB66055 and AAB15886 to AAB16040 represent
 CC nucleotide and protein sequences associated with E. coli proliferation.
 CC AAB66056 and AAB66057 represent primers used for sequencing E. coli
 CC proliferation inhibiting nucleotide inserts in an example from the
 CC present invention. Methods from the present invention can be used to
 CC identify a proliferation- required gene in a microorganism, by contacting
 CC a microorganism with a proliferation-required gene activity inhibitory

XX Vaccine includes complex comprising FimC-FimH, or PapD-PapG in a carrier.
XX
XX
XX
XX Example 12; Col 25-28; 19pp; English.
XX
XX The invention relates to a vaccine comprising a complex comprising FimC-
XX FimH, or PapD-PapG, in a carrier. Also included is a method for
XX protecting against a bacterial infection caused by pilus-bearing
XX bacteria, comprising administering the vaccine to a human at risk. The
XX vaccine is useful for eliciting production of an antibody against the
XX complex when administered to a patient. It is used for protecting against
XX bacterial infection, e.g. urinary tract infection (UTI) e.g. cystitis,
XX pyelonephritis, or a bladder infection. The present sequence represents
XX E. coli adhesin, FimH, used to make a vaccine of the invention
XX
XX Sequence 300 AA;

Query Match 100.0%; Score 821; DB 6; Length 300;
Best Local Similarity 100.0%; Pred. No. 2e-79;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKRVITLFAVLNGSVNWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
DB 1 MKRVITLFAVLNGSVNWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDVTLQGSAYGGVLSNFSGTVKYSYGFPTTSEPRVYNSRT 120
DB 61 TOIFCHNDYPETITDVTLQGSAYGGVLSNFSGTVKYSYGFPTTSEPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQNNY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQNNY 158

RESULT 6
ADC24381
ID ADC24381 standard; protein; 300 AA.
XX
XX ADC24381;
XX
XX 18-DEC-2003 (first entry)
XX
XX E. coli FimH protein.
XX FimH; bacterial chaperone; immunogenic fragment; adhesin protein;
XX vaccine; bacterial infection; FimH; mannose-binding fragment;
XX urinary tract infection; UTI; enterobacterial infection; FimC;
XX immunisation; adhesin/chaperone complex; bladder infection;
XX antibacterial; enzyme.
XX
XX Escherichia coli.
XX
XX Key Location/Qualifiers
XX Misc-difference 16 /note= "Encoded by tgc"
XX
XX Misc-difference 151 /note= "Encoded by agg"
XX
XX Misc-difference 208 /note= "Encoded by cgt"
XX
XX Misc-difference 299 /note= "Encoded by tat"
XX
XX US2003099665-A1.
XX
XX 29-MAY-2003.
XX
XX 06-NOV-2002; 2002US-00288978.
XX
XX 23-APR-1998; 98US-0082824P.
XX
XX 23-APR-1999; 99US-00298494.
XX
XX (LANG/) LANGERMANN S.
XX (HULT/) HULTGREN S J.

PA (PINK/) PINKNER J S.
PA (AUGU/) AUGUSTE C G.
XX Langermann S, Hultgren SJ, Pinkner JS, Auguste CG;
XX
XX WPI; 2003-678130/64.
XX N-PSDB; ADC24380.
XX
XX New vaccine comprising bacterial chaperone protein or bacterial adhesin
XX protein FimH or a mannose-binding fragment of FimH, useful for preventing
XX or treating enterobacterial infections.
XX
XX Example 12; SEQ ID NO 2; 21pp; English.
XX
XX The invention discloses a vaccine against bacterial infections comprising
XX a complex of a bacterial chaperone protein with an adhesin protein or an
XX immunogenic fragment of the adhesin protein. Also claimed is a vaccine
XX against bacterial infections comprising the bacterial adhesin protein
XX FimH, or a mannose-binding fragment of FimH, an antibody raised against a
XX complex of a bacterial chaperone protein with either an adhesin or an
XX immunogenic fragment of the adhesin protein, an antibody raised against
XX either the bacterial adhesin protein FimH or an immunogenic mannose-
XX binding fragment of FimH, preventing or treating urinary tract infection
XX (UTI) in a host, preventing or treating enterobacterial infections, e.g.
XX host and expressing a protein complex in a bacterial cell. The chaperone
XX protein is FimC. Preventing or treating enterobacterial infections in a
XX UTI in a host comprises immunising the host with the vaccine and at least
XX one antibody raised against a complex of a FimC with either FimH or an
XX immunogenic mannose-binding fragment of FimH. The protein complex is an
XX adhesin/chaperone complex. The vaccine is useful for preventing or
XX treating enterobacterial infections, particularly, the vaccine is used
XX for urinary tract or bladder infections caused by Escherichia coli. The
XX antibody raised against a complex of a bacterial chaperone protein is
XX useful for detecting urinary tract infections and for the prevention
XX and/or treatment of urinary tract infections caused by Escherichia coli.
XX The sequence presented is the E. coli FimH protein.
XX
XX Sequence 300 AA;

Query Match 100.0%; Score 821; DB 7; Length 300;
Best Local Similarity 100.0%; Pred. No. 2e-79; 0; Indels 0; Gaps 0;
Matches 158; Conservative 0; Mismatches 0;
QY 1 MKRVITLFAVLNGSVNWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
DB 1 MKRVITLFAVLNGSVNWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDVTLQGSAYGGVLSNFSGTVKYSYGFPTTSEPRVYNSRT 120
DB 61 TOIFCHNDYPETITDVTLQGSAYGGVLSNFSGTVKYSYGFPTTSEPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQNNY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQNNY 158

RESULT 7
ADG17420
ID ADG17420 standard; protein; 300 AA.
XX
XX ADG17420;
XX
XX 26-FEB-2004 (first entry)
XX
XX Escherichia coli FimH D-mannose-binding adhesin wild-type protein.
XX
XX urinary tract infection; UTI; FimH; antibacterial; virucide; bacterial;
XX viral infection; vaccine; FimH; D-mannose-binding adhesin; wild-type;
XX protein co-ordinate data.
XX
XX Escherichia coli.
XX
XX WO2002102974-A2.
XX

```
XX PD 27-DEC-2002.
XX XX
XX PF 10-DEC-2001; 2001WO-US047994.
XX XX
XX PR 08-DEC-2000; 2000US-0254353P.
XX PR 29-JUN-2001; 2001US-0301878P.
XX XX
XX PA (MEDI-) MEDIMUNE INC.
XX XX
XX PI Langermann SR, Hultgren SJ, Hung C, Bouckaert J;
XX PI WPI; 2003-167503/16.
XX DR N-PSDB; ADG17419.
XX DR
XX XX Preventing, treating or ameliorating one or more symptoms of bacterial or
XX PT viral infection, particularly URI, using immunospecific antibodies that
XX PT bind to antigens of mutant FimH protein.
XX XX
XX PS Claim 53; SEQ ID NO 4; 1:194pp; English.
XX XX
XX CC The invention relates to a novel method for preventing, treating or
XX CC ameliorating one or more symptoms associated with a urinary tract
XX CC infection (URI) in a human subject infected with Escherichia coli
XX CC comprising administering one or more antibodies that immunospecifically
XX CC bind to one or more antigens of a mutant FimH protein having one or more
XX CC amino acid substitutions. The method of the invention has antibacterial
XX CC and virucide applications and may be useful for the prevention, treatment
XX CC and/or amelioration of a bacterial or viral infection, in particular a
XX CC urinary tract infection, via the production of a vaccine. The current
XX CC sequence is that of the Escherichia coli FimH D-mannose-binding adhesin
XX CC wild-type protein of the invention.
XX XX
XX SQ Sequence 300 AA;
XX
XX Query Match 100.0%; Score 821; DB 7; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 2e-79;
XX Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
XX Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
XX
XX QY 61 TQIFCHNDYPETITDYVTLQGSAYGVLNFSFGTVKYGSSYPPTTSETPRVYNSRT 120
XX Db 61 TQIFCHNDYPETITDYVTLQGSAYGVLNFSFGTVKYGSSYPPTTSETPRVYNSRT 120
XX
XX QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
XX Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
XX
XX RESULT 8
XX AAR76774
XX ID AAR76774 standard; protein; 300 AA.
XX AC AAR76774;
XX XX
XX DT 16-OCT-2003 (revised)
XX DT 15-MAR-1996 (first entry)
XX XX
XX DE FimH protein derived from E. coli clinical isolate MJH31-3.
XX XX
XX KW FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
XX KW FimA; FimF; FimG; receptor binding site.
XX XX
XX OS Escherichia coli; clinical isolate MJH31-3.
XX XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..21 /note= "Signal peptide"
XX FT Protein 22..300 /note= "Mature FimH"
```

```
XX PN WO9520657-A1.
XX XX
XX PD 03-AUG-1995.
XX XX
XX PF 27-JAN-1995; 9SWO-DK000042.
XX XX
XX PR 27-JAN-1994; 94US-00187166.
XX XX
XX PA (GXBI-) GX BIOSYSTEMS AS.
XX XX
XX PI Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX PI WPI; 1995-275442/36.
XX DR N-PSDB; AAQ93068.
XX DR
XX XX Receptor specific bacterial adhesins - useful for targeting active
XX PT compounds and microbial cells to locations of receptors.
XX XX
XX PS Example 1; Page 44-45; 152pp; English.
XX XX
XX CC The sequences given in AAR76763-76 are FimH proteins from various E. coli
XX CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
XX CC also intercalated at intervals in the fimbrial organelle. Most forms of
XX CC the FimH adhesin target to, and bind to, oligosaccharide structures
XX CC containing terminally located alpha-D-mannoside residues. FimH contains 4
XX CC cysteine residues assumed to direct folding of the molecule into distinct
XX CC functional domains. For comparison FimA and the minor components FimF and
XX CC FimG only have 2 cysteine residues. The localisation of the cysteine
XX CC residues in FimH points to a tandem arrangement of two ancestral genes.
XX CC Similar amino acids can be found in similar positions in the two halves
XX CC of the FimH protein. The "midway" point is located roughly around residue
XX CC 150 in the mature protein. The two halves or domains of FimH have evolved
XX CC differently with the N-terminal section becoming the domain harbouring
XX CC the receptor binding site, whereas the C-terminal sector became the
XX CC domain of the molecule required for integration into the fimbrial
XX CC organelle. These sequences may be used in the production of a variant
XX CC FimH adhesin which may be useful for targeting active compounds and
XX CC microbial cells to locations comprising selected receptors to which the
XX CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
XX XX
XX SQ Sequence 300 AA;
XX
XX Query Match 99.5%; Score 817; DB 2; Length 300;
XX Best Local Similarity 99.4%; Pred. No. 5.4e-73;
XX Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
XX Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
XX
XX QY 61 TQIFCHNDYPETITDYVTLQGSAYGVLNFSFGTVKYGSSYPPTTSETPRVYNSRT 120
XX Db 61 TQIFCHNDYPETITDYVTLQGSAYGVLNFSFGTVKYGSSYPPTTSETPRVYNSRT 120
XX
XX QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
XX Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
XX
XX RESULT 9
XX AAR76771
XX ID AAR76771 standard; protein; 300 AA.
XX AC AAR76771;
XX XX
XX DT 16-OCT-2003 (revised)
XX DT 15-MAR-1996 (first entry)
XX XX
XX DE FimH protein derived from E. coli clinical isolate KS-54.
XX XX
XX KW FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
XX KW FimA; FimF; FimG; receptor binding site.
```

```

XX OS Escherichia coli; clinical isolate KS-54.
XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT Protein /note= "signal peptide"
XX FT Protein 22..300
XX FT Protein /note= "Mature FimH"
XX PN WO9520657-A1.
XX PD 03-AUG-1995.
XX PF 27-JAN-1995; 95WO-DK000042.
XX PR 27-JAN-1994; 94US-00187166.
XX PA (GXBI-) GX BIOSYSTEMS AS.
XX PI Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX DR WPI: 1995-275442/36.
XX DR N-PSDB; AAQ93073.
XX PT Receptor specific bacterial adhesins - useful for targetting active
XX PT compounds and microbial cells to locations of receptors.
XX PS Example 1; Page 44-45; 152pp; English.
XX CC The sequences given in AAR76763-76 are FimH proteins from various E. coli
XX CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
XX CC also intercalated at intervals in the fimbrial organelle. Most forms of
XX CC the FimH adhesin target to, and bind to, oligosaccharide structures
XX CC containing terminally located alpha-D-mannoside residues. FimH contains 4
XX CC cysteine residues assumed to direct folding of the molecule into distinct
XX CC functional domains. For comparison FimA and the minor components FimF and
XX CC FimG only have 2 cysteine residues. The localisation of the cysteine
XX CC residues in FimH points to a tandem arrangement of two ancestral genes.
XX CC Similar amino acids can be found in similar positions in the two halves
XX CC of the FimH protein. The "midway" point is located roughly around residue
XX CC 150 in the mature protein. The two halves or domains of FimH have evolved
XX CC differently with the N-terminal section becoming the domain harbouring
XX CC the receptor binding site, whereas the C-terminal sector became the
XX CC domain of the molecule required for integration into the fimbrial
XX CC organelle. These sequences may be used in the production of a variant
XX CC FimH adhesin which may be useful for targetting active compounds and
XX CC microbial cells to locations comprising selected receptors to which the
XX CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
XX SQ Sequence 300 AA;

Query Match 99.5%; Score 817; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 5.4e-79;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLMLGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNQLVVDLS 60
DB 1 MKRVITLFAVLMLGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNQLVVDLS 60
QY 61 TQIFCHNDYPTETIDYVTLQGSAYGCVLSNFSGTVKYSGSSYPFPTTSETPRVYNSRT 120
DB 61 TQIFCHNDYPTETIDYVTLQGSAYGCVLSNFSGTVKYSGSSYPFPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLLRQTNNY 158
DB 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLLRQTNNY 158

RESULT 10
AAR76763
ID AAR76763 standard; protein; 300 AA.
XX AC AAR76763;

```

```

XX DT 16-OCT-2003 (revised)
XX DT 15-MAR-1996 (first entry)
XX DE FimH protein derived from E. coli clinical isolate KB21.
XX FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
XX FimA; FimF; FimG; receptor binding site.
XX OS Escherichia coli; clinical isolate KB21.
XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT Protein /note= "signal peptide"
XX FT Protein 22..300
XX FT Protein /note= "Mature FimH"
XX PN WO9520657-A1.
XX PD 03-AUG-1995.
XX PF 27-JAN-1995; 95WO-DK000042.
XX PR 27-JAN-1994; 94US-00187166.
XX PA (GXBI-) GX BIOSYSTEMS AS.
XX PI Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX DR WPI: 1995-275442/36.
XX PT Receptor specific bacterial adhesins - useful for targetting active
XX PT compounds and microbial cells to locations of receptors.
XX PS Example 1; Page 44-45; 152pp; English.
XX CC The sequences given in AAR76763-76 are FimH proteins from various E. coli
XX CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
XX CC also intercalated at intervals in the fimbrial organelle. Most forms of
XX CC the FimH adhesin target to, and bind to, oligosaccharide structures
XX CC containing terminally located alpha-D-mannoside residues. FimH contains 4
XX CC cysteine residues assumed to direct folding of the molecule into distinct
XX CC functional domains. For comparison FimA and the minor components FimF and
XX CC FimG only have 2 cysteine residues. The localisation of the cysteine
XX CC residues in FimH points to a tandem arrangement of two ancestral genes.
XX CC Similar amino acids can be found in similar positions in the two halves
XX CC of the FimH protein. The "midway" point is located roughly around residue
XX CC 150 in the mature protein. The two halves or domains of FimH have evolved
XX CC differently with the N-terminal section becoming the domain harbouring
XX CC the receptor binding site, whereas the C-terminal sector became the
XX CC domain of the molecule required for integration into the fimbrial
XX CC organelle. These sequences may be used in the production of a variant
XX CC FimH adhesin which may be useful for targetting active compounds and
XX CC microbial cells to locations comprising selected receptors to which the
XX CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
XX SQ Sequence 300 AA;

Query Match 99.5%; Score 817; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 5.4e-79;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLMLGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNQLVVDLS 60
DB 1 MKRVITLFAVLMLGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNQLVVDLS 60
QY 61 TQIFCHNDYPTETIDYVTLQGSAYGCVLSNFSGTVKYSGSSYPFPTTSETPRVYNSRT 120
DB 61 TQIFCHNDYPTETIDYVTLQGSAYGCVLSNFSGTVKYSGSSYPFPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLLRQTNNY 158
DB 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLLRQTNNY 158

```



```

RESULT 11
AAR76773
ID AAR76773 standard; protein; 300 AA.
AC AAR76773;
XX
XX
DT 16-OCT-2003 (revised)
DT 15-MAR-1996 (first entry)
XX
XX
DE FimH protein derived from E. coli clinical isolate MJ#9-3.
XX
XX FimH; type 1 fimbriae; organella; adhesin; alpha-D-mannoside residue;
KW FimA; FimF; FimG; receptor binding site.
XX
XX Escherichia coli; clinical isolate MJ#9-3.
OS
XX
XX Key a Location/Qualifiers
XX Peptide 1..21
XX /note= "Signal peptide"
XX Protein 22..300
XX /note= "Mature FimH"
XX
XX WO9520657-A1.
XX
XX 03-AUG-1995.
XX
XX 27-JAN-1995; 95WO-DK000042.
XX
XX 27-JAN-1994; 94US-00187166.
XX
XX (GXBI-) GX BIOSYSTEMS AS.
XX
XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX WPI; 1995-275442/36.
XX N-PSDB; AAQ93067.
XX
XX Receptor specific bacterial adhesins - useful for targetting active
XX compounds and microbial cells to locations of receptors.
XX
XX Example 1; Page 44-45; 152pp; English.
XX
XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
XX clinical isolates. FimH is located at the tip of the type 1 fimbriae and
XX also intercalated at intervals in the fimbrial organelle. Most forms of
XX the FimH adhesin target to, and bind to, oligosaccharide structures
XX containing terminally located alpha-D-mannoside residues. FimH contains 4
XX cysteine residues assumed to direct folding of the molecule into distinct
XX functional domains. For comparison FimA and the minor components FimF and
XX FimG only have 2 cysteine residues. The localisation of the cysteine
XX residues in FimH points to a tandem arrangement of two ancestral genes.
XX Similar amino acids can be found in similar positions in the two halves
XX of the FimH protein. The "midway" point is located roughly around residue
XX 150 in the mature protein. The two halves or domains of FimH have evolved
XX differently with the N-terminal section becoming the domain harbouring
XX the receptor binding site, whereas the C-terminal sector became the
XX domain of the molecule required for integration into the fimbrial
XX organelle. These sequences may be used in the production of a variant
XX FimH adhesin which may be useful for targetting active compounds and
XX microbial cells to locations comprising selected receptors to which the
XX adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 300 AA;
XX
XX Query Match 99.5%; Score 817; DB 2; Length 300;
XX Best Local Similarity 99.4%; Pred. No. 5.4e-79;
XX Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MKRVITLFAVLMLGMSVNAWSPACKTANGTAIPIGGGSANVYVNLAPVAVNGQNLVVDLS 60
XX
XX 1 MKRVITLFAVLMLGMSVNAWSPACKTANGTAIPIGGGSANVYVNLAPVAVNGQNLVVDLS 60
XX

```

```

QY 61 TOIFCHNDYPETITDYVTLQSGSAYGGVLSNFSGTVKYSGSYFFPTTSEPRVVNSRT 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 TOIFCHNDYPETITDYVTLQSGSAYGGVLSNFSGTVKYSGSYFFPTTSEPRVVNSRT 120
QY 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILROTNYY 158
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILROTNYY 158

RESULT 12
ADG17467
ID ADG17467 standard; protein; 300 AA.
XX
XX ADG17467;
XX
XX 26-FEB-2004 (first entry)
XX
XX Escherichia coli FimH D-mannose-binding adhesin mutant protein Q154N.
DE
XX urinary tract infection; UTI; FimH; antibacterial; virucide; bacterial;
KW viral infection; vaccine; FimH; D-mannose-binding adhesin; mutant;
KW mutein; protein co-ordinate data.
XX
XX Escherichia coli.
XX
XX Key Location/Qualifiers
XX Misc-difference 154
XX /note= "Wild-type Gln substituted for Asn"
XX
XX WO2002102974-A2.
XX
XX 27-DEC-2002.
XX
XX 10-DEC-2001; 2001WO-US047994.
XX
XX 08-DEC-2000; 2000US-0254353P.
XX 29-JUN-2001; 2001US-0301878P.
XX
XX (MEDI-) MEDIMMUNE INC.
XX
XX Langermann SR, Hultgren SJ, Hung C, Bouckaert J;
XX WPI; 2003-167503/16.
XX
XX Preventing, treating or ameliorating one or more symptoms of bacterial or
XX viral infection, particularly UTI, using immunospecific antibodies that
XX bind to antigens of mutant FimH protein.
XX
XX Claim 52; Page; 1194pp; English.
XX
XX The invention relates to a novel method for preventing, treating or
XX ameliorating one or more symptoms associated with a urinary tract
XX infection (UTI) in a human subject infected with Escherichia coli
XX comprising administering one or more antibodies that immunospecifically
XX bind to one or more antigens of a mutant FimH protein having one or more
XX amino acid substitutions. The method of the invention has antibacterial
XX and virucide applications and may be useful for the prevention, treatment
XX and/or amelioration of a bacterial or viral infection, in particular a
XX urinary tract infection, via the production of a vaccine. The current
XX sequence is that of the Escherichia coli FimH D-mannose-binding adhesin
XX mutant protein of the invention which contains a Q154N mutation. This
XX sequence is not shown within the specification per se but was created by
XX the indexer used information from SEQ ID 4.
XX
XX Sequence 300 AA;
XX
XX Query Match 99.4%; Score 816; DB 7; Length 300;
XX Best Local Similarity 99.4%; Pred. No. 6.9e-79;
XX Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MKRVITLFAVLMLGMSVNAWSPACKTANGTAIPIGGGSANVYVNLAPVAVNGQNLVVDLS 60
XX
XX 1 MKRVITLFAVLMLGMSVNAWSPACKTANGTAIPIGGGSANVYVNLAPVAVNGQNLVVDLS 60
XX

```

Db 1 MKRVITLFAVLGMGWSNVANSFACKTANGTAIPIGGGSANVYNLAPVNVNQNLVVDLS 60
QY 61 TQIFCHNDYPETITDVTLQGSAYGGVLSNFSGTVKYSGSSYPFPTTSETPRVVNSRT 120
Db 61 TQIFCHNDYPETITDVTLQGSAYGGVLSNFSGTVKYSGSSYPFPTTSETPRVVNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 13
AAR76768
ID AAR76768 standard; protein; 300 AA.
XX
AC AAR76768;
XX
DT 16-OCT-2003 (revised)
DT 15-MAR-1996 (first entry)
XX
DE FimH protein derived from E. coli clinical isolate CSH-50.
XX
DE FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
KW FimA; FimF; FimG; receptor binding site.
XX
OS Escherichia coli; clinical isolate CSH-50.
XX
XX Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..300
FT /note= "Mature FimH"
FT
FT
FN W09520657-A1.
XX
XX 03-AUG-1995.
XX
XX 27-JAN-1995; 95WO-DK000042.
XX
XX 27-JAN-1994; 94US-00187166.
XX
XX (GXBI-) GX BIOSYSTEMS AS.
XX
XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX WPI; 1995-275442/36.
XX
XX Receptor specific bacterial adhesins - useful for targeting active
XX compounds and microbial cells to locations of receptors.
XX
XX Example 1; Page 44-45; 152pp; English.
XX
XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
XX clinical isolates. FimH is located at the tip of the type 1 fimbriae and
XX also intercalated at intervals in the fimbrial organelle. Most forms of
XX the FimH adhesin target to, and bind to, oligosaccharide structures
XX containing terminally located alpha-D-mannoside residues. FimH contains 4
XX cysteine residues assumed to direct folding of the molecule into distinct
XX functional domains. For comparison FimA and the minor components FimF and
XX FimG only have 2 cysteine residues. The localisation of the cysteine
XX residues in FimH points to a tandem arrangement of two ancestral genes.
XX Similar amino acids can be found in similar positions in the two halves
XX of the FimH protein. The "midway" point is located roughly around residue
XX 150 in the mature protein. The two halves or domains of FimH have evolved
XX differently with the N-terminal section becoming the domain harbouring
XX the receptor binding site, whereas the C-terminal sector became the
XX domain of the molecule required for integration into the fimbrial
XX organelle. These sequences may be used in the production of a variant
XX FimH adhesin which may be useful for targeting active compounds and
XX microbial cells to locations comprising selected receptors to which the
XX adhesins bind. (Updated on 18-OCT-2003 to standardise OS field)
XX
XX Sequence 300 AA;

Query Match 99.3%; Score 815; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 8.8e-79;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKRVITLFAVLGMGWSNVANSFACKTANGTAIPIGGGSANVYNLAPVNVNQNLVVDLS 60
Db 1 MKRVITLFAVLGMGWSNVANSFACKTANGTAIPIGGGSANVYNLAPVNVNQNLVVDLS 60
QY 61 TQIFCHNDYPETITDVTLQGSAYGGVLSNFSGTVKYSGSSYPFPTTSETPRVVNSRT 120
Db 61 TQIFCHNDYPETITDVTLQGSAYGGVLSNFSGTVKYSGSSYPFPTTSETPRVVNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 14
AAR76772
ID AAR76772 standard; protein; 300 AA.
XX
AC AAR76772;
XX
DT 16-OCT-2003 (revised)
DT 15-MAR-1996 (first entry)
XX
DE FimH protein derived from E. coli clinical isolate U221-3.
XX
XX FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
KW FimA; FimF; FimG; receptor binding site.
XX
OS Escherichia coli; clinical isolate U221-3.
XX
XX Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..300
FT /note= "Mature FimH"
FT
FT
FN W09520657-A1.
XX
XX 03-AUG-1995.
XX
XX 27-JAN-1995; 95WO-DK000042.
XX
XX 27-JAN-1994; 94US-00187166.
XX
XX (GXBI-) GX BIOSYSTEMS AS.
XX
XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX WPI; 1995-275442/36.
XX
XX N-PSDB; AAQ93074.
XX
XX Receptor specific bacterial adhesins - useful for targeting active
XX compounds and microbial cells to locations of receptors.
XX
XX Example 1; Page 44-45; 152pp; English.
XX
XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
XX clinical isolates. FimH is located at the tip of the type 1 fimbriae and
XX also intercalated at intervals in the fimbrial organelle. Most forms of
XX the FimH adhesin target to, and bind to, oligosaccharide structures
XX containing terminally located alpha-D-mannoside residues. FimH contains 4
XX cysteine residues assumed to direct folding of the molecule into distinct
XX functional domains. For comparison FimA and the minor components FimF and
XX FimG only have 2 cysteine residues. The localisation of the cysteine
XX residues in FimH points to a tandem arrangement of two ancestral genes.
XX Similar amino acids can be found in similar positions in the two halves
XX of the FimH protein. The "midway" point is located roughly around residue
XX 150 in the mature protein. The two halves or domains of FimH have evolved
XX differently with the N-terminal section becoming the domain harbouring

CC the receptor binding site, whereas the C-terminal sector became the
 CC domain of the molecule required for integration into the fimbrial
 CC organelle. These sequences may be used in the production of a variant
 CC FimH adhesin which may be useful for targeting active compounds and
 CC microbial cells to locations comprising selected receptors to which the
 CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 300 AA;

Query Match 98.8%; Score 811; DB 2; Length 300;
 Best Local Similarity 98.7%; Pred. No. 2.4e-78;
 Matches 156; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
 DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
 QY 61 TQIFCHNDYPETITDYVTLQGSAYGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPETITDYVTLQGSAYGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRT 120
 QY 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
 DB 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 15

AAR76770
 ID AAR76770 standard; protein; 300 AA.

XX AAR76770;
 AC
 XX
 DT 16-OCT-2003 (revised)
 DT 15-MAR-1996 (first entry)
 XX
 XX FimH protein derived from E. coli clinical isolate F-18.
 DE FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
 KW FimH; FimF; FimG; receptor binding site.
 XX
 OS Escherichia coli; clinical isolate F-18.
 XX

FH Key Location/Qualifiers
 FT Peptide 1..21 /note= "Signal peptide"
 FT Protein 22..300 /note= "Mature FimH"
 FT
 XX

PN WO9520657-A1.

XX
 XX 03-AUG-1995.
 XX
 XX 27-JAN-1995; 95WO-DK000042.
 XX
 XX 27-JAN-1994; 94US-00187166.
 XX
 XX (GXBI-) GX BIOSYSTEMS AS.

XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
 XX
 XX WPI: 1995-275442/36.
 XX N-PSDB; AAQ93063.

XX Receptor specific bacterial adhesins - useful for targeting active
 PT compounds and microbial cells to locations of receptors.
 XX
 XX Example_1; Page 44-45; 152pp; English.

XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
 CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
 CC also intercalated at intervals in the fimbrial organelle. Most forms of
 CC the FimH adhesin target to, and bind to, oligosaccharide structures
 CC containing terminally located alpha-D-mannoside residues. FimH contains 4

CC cysteine residues assumed to direct folding of the molecule into distinct
 CC functional domains. For comparison FimA and the minor components FimF and
 CC FimG only have 2 cysteine residues. The localisation of the cysteine
 CC residues in FimH points to a tandem arrangement of two ancestral genes.
 CC Similar amino acids can be found in similar positions in the two halves
 CC of the FimH protein. The "midway" point is located roughly around residue
 CC 150 in the mature protein. The two halves or domains of FimH have evolved
 CC differently with the N-terminal section becoming the domain harbouring
 CC the receptor binding site, whereas the C-terminal sector became the
 CC domain of the molecule required for integration into the fimbrial
 CC organelle. These sequences may be used in the production of a variant
 CC FimH adhesin which may be useful for targeting active compounds and
 CC microbial cells to locations comprising selected receptors to which the
 CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 300 AA;

Query Match 98.5%; Score 809; DB 2; Length 300;
 Best Local Similarity 98.1%; Pred. No. 3.9e-78;
 Matches 155; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
 DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
 QY 61 TQIFCHNDYPETITDYVTLQGSAYGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPETITDYVTLQGSAYGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRT 120
 QY 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
 DB 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

Search completed: November 27, 2004, 16:13:23
 Job time : 155 secs

